Formalizing Factorization on Euclidean Domains and Abstract Euclidean Algorithms [⋆]

Thaynara Arielly de Lima[1⋆⋆], Andréia Borges Avelar[3], André Luiz Galdino[2], and Mauricio Ayala-Rincón[3⋆⋆⋆]

[1] Universidade Federal de Goiás, Brasil
[2] Universidade Federal de Catalão, Brasil
[3] Universidade de Brasília, Brasil

Abstract. This paper discusses the extension of the PVS sub theory for rings, part of the PVS `algebra` theory, with theorems related to the division algorithm for Euclidean rings and Unique Factorization Domains that are general structures where an analogous of the Fundamental Theorem of Arithmetic holds. First, we formalize the general abstract notions of divisibility, prime and irreducible elements in commutative rings, which are essential to dealing with unique factorization domains. Then, we formalize the landmark theorem that establishes that every principal ideal domain is a unique factorization domain. Finally, we specify the theory of Euclidean domains and formally verify that the rings of integers, the Gaussian integers, and arbitrary fields are Euclidean domains. To highlight the benefits of such a general abstract discipline of formalization, we specify a Euclidean gcd algorithm for Euclidean domains and formalize its correctness. Also, we show how this correctness is inherited under adequate parameterizations for the structures of integers and Gaussian integers.

Keywords: Theorem Proving · Proof Assistants · PVS · Unique Factorization Domains · Euclidean Rings · Division Algorithms.

1 Introduction

The NASA PVS `algebra` library ([3]) was recently enriched with a series of theorems related to the theory of rings. The extension includes complete formalizations of the isomorphism theorems for rings, principal and prime and maximal ideals, and a general abstract version of the Chinese Remainder Theorem (CRT) which holds for abstract rings, including non-commutative rings. The benefit of formalizing algebraic results from this abstract theoretical perspective was made evident showing how, from the abstract version of CRT, the well-known numerical version of CRT for the ring of integers $\mathbb{Z}$ was formalized [18].

[⋆] Supported by FAPDF DE 00193.00001175/21-11 and CNPq Universal 409003/21-2 grants. Fourth author partially funded by CNPq grant 313290/21-0.

[⋆⋆] ✉ Thaynara Arielly de Lima `thaynaradelima@ufg.br`.

[⋆⋆⋆] ✉ Mauricio Ayala-Rincón `ayala@unb.br`.

In this work, we give another substantial step towards enriching the PVS abstract algebra library by formalizing properties about factorization in commutative rings regarding both unique factorization domains and Euclidean rings. Roughly, unique factorization domains are abstract structures for which a general version of the Fundamental Theorem of Arithmetic holds. On the other hand, Euclidean rings are equipped with a norm that allows defining a suitable generalization of Euclid's division lemma and consequently of notions as for example the one of greatest common divisor (`gcd`). The practicality of `gcd` is well-known in the ring $\mathbb{Z}$. Nevertheless, mathematicians known this notion is of general fundamental importance in abstract Euclidean domains for which in general, `gcd` should and may be defined in different manners.

Figure 1 highlights the subtheories subject of the extension to the PVS theory `algebra` discussed in this paper. The red ones are related to Euclidean rings and `gcd` algorithms for Euclidean domains, and the orange ones are those related to unique factorization domains. The extension includes 210 new formulas enlarging the theory `algebra` from 1356 (cf [18]) to 1566 formalized lemmas.

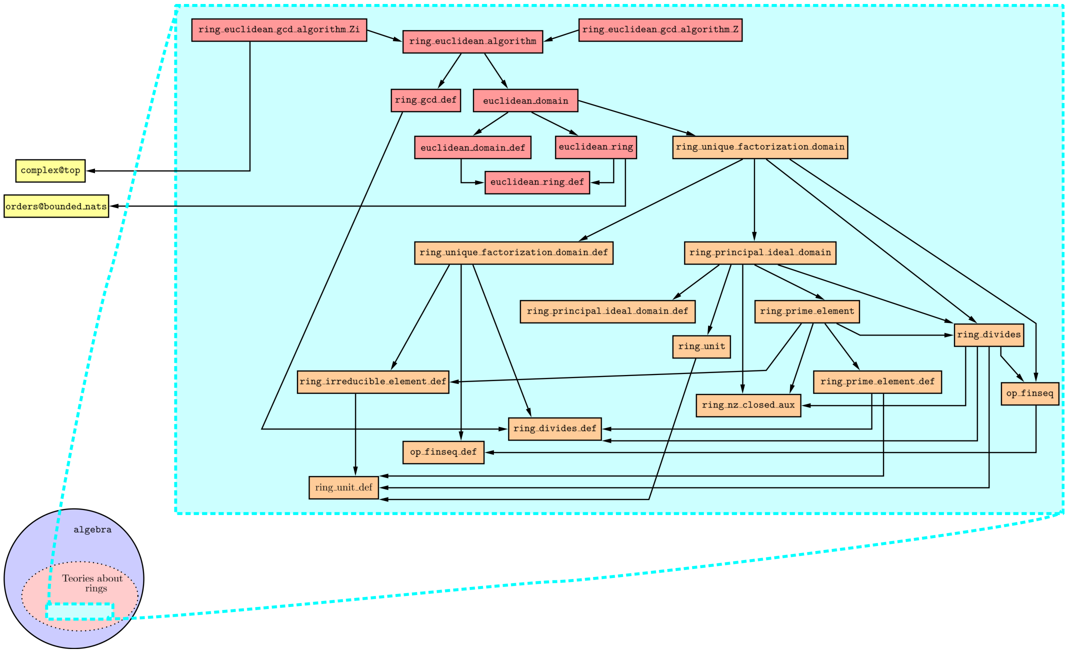

Fig. 1. Ring theories expanding the PVS `algebra` library

The main motivation to formalize such structures is due to their potential theoretical and practical applications. Using the example of `gcd`, one can provide a general abstract version of the Euclidean algorithm to determine a `gcd` between two elements (Euclidean `gcd` algorithm) in a Euclidean domain. Since the ring of integers $\mathbb{Z}$, the Gaussian integers $\mathbb{Z}[i]$ (which are the subset of complex numbers whose real and imaginary parts are integer numbers) and rings of polynomials over integral domains are particular Euclidean domain structures, the Euclidean `gcd` algorithm can be applied over them, in a relatively straightforward manner, to compute `gcd`s in different manners. Not only for the above mentioned structures, but for a variety of Euclidean domains.

Also, every element of a unique factorization domain can be factorized as a finite number of irreducible elements, and one can prove that Euclidean domains are unique factorization domains. These properties allow us to introduce modular arithmetic, and verify generic versions of Euler's Theorem and Fermat's Little Theorem for Euclidean domains, and promote factorization in Euclidean domains as a convenient feature to develop efficient algorithms in symbolic computation [17], [9]. Thus, a formalization of the main results about unique factorization and Euclidean domains would allow the formal verification of more complex theories involving such structures in their scope.

The main contributions of this paper are listed below.

- We formalize the abstract notions of divisibility, prime and irreducible elements in commutative rings, which are essential to deal with unique factorization domains. In integral domains, prime elements are irreducible. The converse is not true in general. Among other properties, we formalize the theorem that establishes that in principal ideal domains (as it is well-known, it holds in $\mathbb{Z}$) irreducible elements are also prime.
- We specify unique factorization domains and formalize the theorem that every principal ideal domain is a unique factorization domain, which is a landmark result in abstract algebra.
- We specify the notion of Euclidean domain and formally verify that the rings $\mathbb{Z}$ and $\mathbb{Z}[i]$, and any arbitrary field are Euclidean domains.
- We specify the general abstract notion of `gcd` for commutative rings, providing a general Euclidean `gcd` algorithm for Euclidean domains and formalize its correctness. Using this result, we parameterize the adequate norms and `gcd` relations for the rings $\mathbb{Z}$ and $\mathbb{Z}[i]$; thus, obtaining in a straightforward manner the correctness of such instantiations of the abstract algorithm for these Euclidean domains. In this manner, we illustrate the benefits of maintaining the abstract general discipline of formalization for algebraic theories and the potential of such a discipline for application in concrete algebraic structures.

The paper is organized as follows: Section 2 presents a theoretical overview of unique factorization and Euclidean domains, pointing out the main concepts and results. Also, it comments on some differences between pen-and-paper proofs presented in Hungerford's textbook [14] and this formalization. Section 3 discusses the aspects of the formalization of the Euclidean `gcd` Algorithm for Euclidean Domains, as well as its application for two particular cases. Section 4 brings related work. Finally, Section 5 concludes and suggests future work. The formalizations were developed using the Prototype Verification System (PVS) and are available at algebra.

2 Formalization of Euclidean Domains

Notions such as prime element, division, and `gcd` between two elements and some landmark results, including the Fundamental Theorem of Arithmetic, Euclid's division lemma, and Euclidean Algorithm, are well established and widespread for the ring of integers. Such concepts and general versions of exciting results are extended for abstract algebraic structures [14],[8] and are the scope of our formalization. This section gives both a theoretical overview of the central notions and properties formalized and discuss their formalization. To point out some differences between pen-and-paper vs. formalized proofs, some analytical concepts and results are presented as they are enunciated in Chapter III of Hungerford's textbook [14].

2.1 Prime and irreducible elements on rings

The definitions of prime and irreducible elements rely on the general concept of divisibility on a ring. The specification of the notions of divisibility and associated elements are given in Specification 1.1.

Specification 1.1. Specification of the definition of divisibility and associated elements in the subtheory `ring_divides_def`

```
R: VAR (ring?)
a, b: VAR T

divides?(R)(a: (R - {zero}), b: (R)): bool =  EXISTS (x: (R)): a*x = b

associates?(R)(a,b:(R - {zero})): bool = divides?(R)(a,b) AND
                                         divides?(R)(b,a)
```

In Hungerford's textbook the definition of divisibility relies on a commutative ring. In fact, it avoids the discrimination between left or right divisor of an element, and since the main results demand a commutative ring in the hypothesis, it is a reasonable requirement. However, notice that commutativity is not a crucial property in such a notion, since it only depends on the operation of multiplication in a ring. Because of that, we opted for generalize the definition and specify divisibility on rings, not necessarily commutative. Another remark is about the specification of `associates?(R)(a,b)`: the textbook omits that `a` and `b` are nonzero elements. Of course, this is obvious since it is required in the definition of `divides?(R)(a,b)`. However, the lack of such a hypothesis is recurrent in several statements along the text that require it (for example, in Theorem 1).

In the subtheory `ring_divides` , we formalized the properties related to the divisibility stated in Theorem 1. Some of them involve the object "unit". In a ring $(R, +, *, zero, one)$ with identity one for multiplication, an element u is called a *unit* if u is left- and right-invertible; that is, there exist elements $u_1^{-1}, u_2^{-1} \in R$ such that $u * u_1^{-1} = u_2^{-1} * u = one$.

Theorem 1 (Theo. 3.2 - Hungerford). *Let a, b and u be elements of a commutative ring R with identity.*

i) *a divides b (denoted as $a \mid b$) if and only if $(b) \subset (a)$, where (x) denotes the principal ideal generated by x.*
ii) *a and b are associates if and only if $(a) = (b)$.*
iii) *u is a unit if and only if $u \mid r$ for all $r \in R$.*
iv) *u is a unit if and only if $(u) = R$.*
v) *The relation "a and b are associates" is an equivalence relation on R.*
vi) *If $a = br$, where $r \in R$ is a unit, then a and b are associates. If R is an integral domain, then the converse is true.*

We specified prime and irreducible elements on a ring with identity (Specification 1.2) from the concepts of divisibility and unit.

Specification 1.2. Specification of irreducible and prime elements in the subtheories `ring_irreducible_element_def` and `ring_prime_element_def` , respectively

```
R: VAR (ring_with_one?)
R_irreducible_element?(R)(x:(R)): bool = x/=zero AND (NOT unit?(R)(x)) AND
    (FORALL (a,b:(R)): x = a*b IMPLIES (unit?(R)(a) OR unit?(R)(b)))
%-------------------------------------
R_prime_element?(R)(x:(R)): bool = x/=zero AND (NOT unit?(R)(x)) AND
 (FORALL (a,b:(R)): divides?(R)(x, a*b) IMPLIES
                    divides?(R)(x, a) OR divides?(R)(x, b))
```

The ring of integers has the feature that prime and irreducible elements are indistinguishable. However, this is not true for some algebraic structures. For instance, 2 is prime but not irreducible in $\mathbb{Z}_6$. Theorem 2 gives some properties regarding prime and irreducible elements formalized in the subtheory `ring_prime_element` . Among others, it provides the result that prime and irreducible elements are equal over principal ideal domains.

Theorem 2 (Theo. 3.4 - Hungerford). *Let p and c be nonzero elements is an integral domain R.*

i) *p is prime if and only if (p) is nonzero prime ideal;*
ii) *c is irreducible if and only if (c) is maximal in the set S of all proper principal ideals of R.*
iii) *Every prime element of R is irreducible.*
iv) *If R is a principal ideal domain, then p is prime if and only if p is irreducible.*
v) *Every associate of an irreducible [resp. prime] element of R is irreducible [resp. prime].*
vi) *The only divisors of an irreducible element of R are its associates and the units of R.*

Although the result is stated for integral domains, Hungerford advises us that, in some parts of the theorem, weakened hypothesis can be considered. We formalize the results using the minimum number of required conditions and detect that items (i) and (vi) of the Theorem 2 hold for commutative rings with identity.

2.2 Unique Factorization Domains

The well-known Fundamental Theorem of Arithmetic for integers states that any positive integer number greater than 1 can be factorized as a unique product of prime numbers, unless a permutation of such factors. Unique Factorization Domains (UFDs) are integral domains where an analogous of such a Theorem holds. The Specification 1.3 shows the definition of UFDs in PVS. It depends on a sequence of irreducible elements `fsIr?`$(R)(fsI)$ on a ring R with identity and a recursive operator `op_fseq`(fsI), specified in subtheory `op_finseq_def` , which multiplies the elements of such a sequence. We specify the operator `op_fseq`(fsI) over an abstract structure $(T, *, one)$ equipped with a binary operation $*$ and a constant *one*. From the point of view of the formalization, such a general specification is very useful for two reasons: it allows the use of the operator `op_fseq`(fsI) in a variety of abstract and concrete structures (monoids, monads, groups, rings, integers, reals) by only adequately parameterizing the subtheory `op_finseq_def`; also, it avoids proof obligations called *Type Correctness Conditions (TCCs)* generated by the system, since the operator is defined for elements of an abstract type, which provides more automation in our formal verification.

Specification 1.3. Specification of Unique Factorization subtheory `ring_unique_factorization_domain_def`

```
fsIr?(R)(fsI: finseq[(R)]): bool = FORALL (i: below[length(fsI)]):
     R_irreducible_element?(R)(fsI(i))

unique_factorization_domain?(R): bool = integral_domain_w_one?(R) AND
FORALL(a: (R)): a /= zero AND NOT unit?(R)(a) IMPLIES
 EXISTS(fsI:(fsIr?(R))):a = op_fseq(fsI) AND
 FORALL(fsIp:fsIr(R)):a = op_fseq(fsIp) IMPLIES length(fsI) = length(fsIp) AND
 EXISTS(phi:[below[length(fsI)]->below[length(fsI)]]): (bijective?(phi)) AND
 FORALL(i:below[length(fsI)]): associates?(R)(fsIp(phi(i)),fsI(i))
```

In subtheory `ring_unique_factorization_domain`, we formalized the Theorem 3, which is a landmark result about UFDs.

Theorem 3 (Theo. 3.7 - Hungerford). *Every principal ideal domain R is a unique factorization domain.*

The formalization of the Theorem 3 has two main steps. We briefly comment on the following.

Step 1 - Existence of a factorization

First, we enriched previous subtheories established in the theory `algebra` with auxiliary results. In the subtheory `ring_ideal` , we formalized the lemma `chain_ideal_union ideal` that states that the union of a chain of ideals in a ring R is an ideal. In the subtheory `ring_with_one_maximal_ideal` , we formalized the lemma `nonzero_ring_exists_maximal_ideal_aux` , which proves that every ideal in a ring R with identity, except R itself, is contained in a maximal ideal in R. The formalization of this lemma considers an ideal $A \neq R$, $S = \{B \subset R;\ B$ is ideal in $R,\ B \neq R$ and $A \subset B\}$ and $\mathcal{C} = \{C_i | i \in I\}$

an arbitrary chain of ideals in S. We prove that the ideal $C = \bigcup C_i$ is an upper bound of the chain $\mathcal{C}$ in S and, by using Zorn's lemma from NASA theory `orders`, we conclude that S has a maximal element, which is a maximal ideal in R. In the subtheory `ring_principal_ideal` , we add the lemma `stable_chain` , which says that if R is a principal ideal ring and $(a_1) \subset (a_2) \ldots$ is a chain of ideals in R, then for some positive integer n, $(a_j) = (a_n)$ for all $j \geq n$. The lemma `nonzero_nonunit_irreducible_divides` , formalized in the subtheory `ring_principal_ideal_domain` , states that every nonzero and nonunit element in a principal ideal domain is divided by an irreducible element.

We conclude Step 1 by verifying that the subset below, of R, a principal ideal domain, is empty.

$$\texttt{non\_fact\_el\_set}(R) = \{\ x : x \text{ is a nonzero nonunit element in } R \text{ and cannot be finitely factorized into irreducible elements.}\}$$

In fact, if $a \in$ `non_fact_el_set`(R), we could build an ascending chain $(a) \subset (a_1) \subset \ldots$ of ideals, which is not possible by the lemma `stable_chain`. The key to verify such fact was to specify the recursive function `phi`(n, R, a) showed in Specification 1.4 (subtheory `ring_principal_ideal_domain`) and verify that it is well defined whenever `non_fact_el_set`(R) is nonempty.

Specification 1.4. Auxiliary function to build an ascending chain of ideals

```
phi(n:nat, R:principal_ideal_domain, a:(non_fact_el_set(R))):
 RECURSIVE (non_fact_el_set(R)) =
  IF n = 0 then a
  ELSE  choose ({x : (non_fact_el_set(R))|
                strict_subset?(one_gen(R)(phi(n-1, R, a)),one_gen(R)(x))})
                ENDIF  MEASURE n
```

If $a \in$ `non_fact_el_set`(R), the choice of the element a_1, obtained by the function `choose` in Specification 1.4, is guaranteed. In fact, the lemma `nonzero_nonunit_irreducible_divides` ensures that $a = ca_1$, where c is irreducible. It implies that a_1 belongs to `non_fact_el_set`(R) and satisfies the condition $(a) \subset (a_1)$ by Theorem 1(i).

Step 2: "Uniqueness" of a factorization

We mean "uniqueness", the existence of a bijective function between the elements of two factorizations mapping associated elements. First, we formalized the lemma `prime_el_divides` (subtheory `ring_prime_element`) which states if a prime element p in an integral domain divides the product $a_1 \ldots a_n$ then there exists $1 \leq i \leq n$ such that p divides a_i. By 2(iii), it holds if p is an irreducible element. From this, if $a_1 \ldots a_n = a = b_1 \ldots b_m$, where $a_i, 1 \leq n$ and b_j, $1 \leq m$ are irreducible elements, then c_1 divides d_i, for some i. By Theorem 2 (vi), c_1 and d_i are associates. Using induction on n, we prove that $n = m$ and establish the required bijective function.

2.3 Euclidean Rings

A Euclidean ring is a commutative ring R equipped with a norm φ over $R - \{zero\}$, where an abstract version of the well-known Euclid's division lemma holds. We specify Euclidean rings and Euclidean domains in the subtheories `euclidean_ring_def` and `euclidean_domain_def` (Specification 1.5).

Specification 1.5. Definitions of Euclidean rings and Euclidean domains

```
euclidean_ring?(R): bool = commutative_ring?(R) AND
EXISTS (phi: [(R - {zero}) -> nat]): FORALL(a,b: (R)):
  ((a*b /= zero IMPLIES phi(a) <= phi(a*b)) AND
   (b /= zero IMPLIES EXISTS(q,r:(R)):
    (a = q*b+r AND (r = zero OR (r /= zero AND phi(r) < phi(b))))))

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

euclidean_domain?(R): bool = euclidean_ring?(R) AND integral_domain_w_one?(R)
```

In subtheory `euclidean_domain` , we formalized that elements of Euclidean ring can be factorized as irreducible elements by verifying Theorem 4.

Theorem 4 (Theo. 3.9 - Hungerford). *A Euclidean ring R is a principal ideal ring with identity. Consequently every Euclidean domain is a unique factorization domain.*

The verification makes use of the well-ordering principle over $\varphi(I^*) = \{\varphi(x) \in \mathbb{N};;\ \mathrm{x} \in I - \{zero\}\}$, where I is a nonzero ideal in R and φ is a norm on $R - \{zero\}$. By choosing $a \in I$ such that $\varphi(a)$ is the minimum element of $\varphi(I^*)$, $b \in I$ satisfies $b = qa + r$, for some $q \in R$ and $r \in I$. From this, we infer that $r = 0$, since $r \neq 0$ contradicts the minimality of $\varphi(a)$. Consequently, $b = qa$ and $I \subset Ra \subset (a) \subset I$, which guarantee that every ideal in R is a principal ideal. By 3, we have that a Euclidean principal ideal domain is a unique factorization domain.

In subtheory `euclidean_domain` , we also formalized the results stating that the ring of integers ( ) and any arbitrary field ( ) are Euclidean domains.

3 Formalization of `gcd` Algorithm for Euclidean Domains

Two additional definitions were included in the theory `Euclidean_ring_def` to allow abstraction of adequate Euclidean norms and associated functions fulfilling the properties of Euclidean rings (see Specification 1.6).

The first definition is the relation `Euclidean_pair?` Given a Euclidean ring R and a Euclidean norm of non zero elements over the naturals $\phi : R \setminus \{zero\} \to \mathbb{N}$, `Euclidean_pair?`$(R, \phi)$ holds whenever ϕ satisfies the constraints of a Euclidean norm over R.

The second definition is the curried relation `Euclidean_f_phi?`$(R, \phi)(f_\phi)$. It holds whenever `Euclidean_pair?`(R, ϕ) holds, and f_ϕ a function from $R \times R \setminus \{zero\}$ to $R \times R$, such that for all pair of elements of R in its domain, $f_\phi(a, b)$ gives a pair of elements, say (div, rem) satisfying the constraints of Euclidean rings regarding the norm ϕ: if $a \neq zero$, $a = div * b + rem$, and if

$rem \neq zero$, $\phi(rem) < \phi(b)$. These definitions are correct since the existence of such a ϕ and f_ϕ is guaranteed by the fact that R is a Euclidean ring. Also, notice that the decrement of the norm, i.e., $\phi(rem) < \phi(b)$, is the key to build an abstract Euclidean terminating procedure.

Specification 1.6. Additional definitions in the subtheory `Euclidean_ring_def`

```
Euclidean_pair?(R : (Euclidean_ring?), phi: [(R - {zero}) -> nat]) : bool =
    FORALL(a,b: (R)): ((a*b /= zero IMPLIES phi(a) <= phi(a*b)) AND
                       (b /= zero IMPLIES
                         EXISTS(q,r:(R)): (a = q*b+r AND
                             (r = zero OR (r /= zero AND phi(r) < phi(b))))))

Euclidean_f_phi?(R : (Euclidean_ring?),
                 phi : [(R - {zero}) -> nat] | Euclidean_pair?(R,phi))
                (f_phi : [(R) , (R - {zero}) -> [(R),(R)]]) : bool =
                 FORALL (a : (R), b :(R - {zero})):
                  IF a = zero THEN f_phi(a,b) = (zero, zero)
                  ELSE LET div = f_phi(a,b)`1, rem = f_phi(a,b)`2 IN
                     a = div * b + rem AND
                    (rem = zero OR (rem /= zero AND phi(rem) < phi(b)))
                  ENDIF
```

Using previous two relations, a general abstract recursive Euclidean `gcd` algorithm is specified as the curried `Euclidean_gcd_algorithm` in the subtheory `ring_euclidean_algorithm` (See Specification 1.7). The correctness of this algorithm is guaranteed by the types of its arguments. Indeed, since allowed arguments R, ϕ, and f_ϕ should satisfy `Euclidean_f_phi?`$(R,\phi)(f_\phi)$, R is a Euclidean ring with associated Euclidean norm ϕ and adequate division and remainder functions given by f_ϕ. The termination of the algorithm is a proof obligation (termination TCC) automatically generated by PVS. Termination is proved using the lexicographic `MEASURE` of the algorithm provided in the specification. This measure decreases after each possible recursive call: for `Euclidean_gcd_algorithm`$(R,\phi,f_\phi)(a,b)$, if $a \neq zero$, $\phi(a) \geq \phi(b)$ and $rem \neq zero$, the recursive call is `Euclidean_gcd_algorithm`$(R,\phi,f_\phi)(b,rem)$; thus, $(\phi(b),\phi(a))$ is lexicographical greater than $(\phi(rem),\phi(b))$, since $\phi(b) > \phi(rem)$. In the other case, if $a \neq zero$, and $\phi(b) > \phi(a)$, the recursive call is `Euclidean_gcd_algorithm`$(R,\phi,f_\phi)(b,a)$; thus, $(\phi(b),\phi(a))$ is lexicographical greater than $(\phi(a),\phi(b))$, since $\phi(b) > \phi(a)$.

The proof of correctness of the recursive algorithm, is given as a straightforward corollary of the `Euclid_theorem` (in Specification 1.7) that establishes the correctness of each recursive step regarding the abstract definition of `gcd` given in Specification 1.8. Essentially, what this theorem states is that, given an adequate Euclidean norm ϕ and associated function f_ϕ, the `gcd` of a pair (a,b) is equal to the `gcd` of the pair (rem,b), where rem is computed through f_ϕ, i.e., rem is equal to the second projection of $f_\phi(a,b)$. Notice, that since Euclidean rings allow a variety of Euclidean norms and associated functions (e.g., [14], [10]), the definition of `gcd` is specified as the relation `gcd`? and not as a function.

Finally, the proof of correctness of the abstract Euclidean algorithm is by induction, using the lexicographic `MEASURE` of the algorithm, as the theorem `Euclidean_gcd_alg_correctness` (in Specification 1.7). For an input pair

(a, b), in the inductive step of the proof, when $\phi(b) > \phi(a)$ and the recursive call swaps the arguments, one assumes that

$$\texttt{gcd?}(R)(\{b, a\}, \texttt{Euclidean\_gcd\_algorithm}(R, \phi, f_\phi)(b, a))$$

which means that `Euclidean_gcd_algorithm`$(R, \phi, f_\phi)(b, a)$ computes correctly the `gcd` of the pair (b, a). From this assumption, one concludes that

$$\texttt{gcd?}(R)(\{a, b\}, \texttt{Euclidean\_gcd\_algorithm}(R, \phi, f_\phi)(a, b))$$

Otherwise, when $\phi(a) \geq \phi(b)$, $rem = (f_\phi(a, b))'2$, and the recursive call is `Euclidean_gcd_algorithm`$(R, \phi, f_\phi)(b, rem)$, by induction hypothesis one has that

$$\texttt{gcd?}(R)(\{b, rem\}, \texttt{Euclidean\_gcd\_algorithm}(R, \phi, f_\phi)(b, rem))$$

Finaly, by application of `Euclid_theorem`, one concludes that the abstract general Euclidean algorithm computes correctly a `gcd` for the pair (a, b).

Specification 1.7. Abstract `gcd` Euclidean algorithm for Euclidean rings in the subtheory `ring_euclidean_algorithm`

```
Euclidean_gcd_algorithm(R : (Euclidean_domain?[T,+,*,zero,one]),
                        (phi: [(R - {zero}) -> nat] | Euclidean_pair?(R,phi)),
                        (f_phi: [(R),(R - {zero}) -> [(R),(R)]] |
                                       Euclidean_f_phi?(R,phi)(f_phi)))
                        (a: (R), b: (R - {zero})) : RECURSIVE (R - {zero}) =
  IF  a = zero THEN b
  ELSIF  phi(a) >= phi(b) THEN
      LET rem = (f_phi(a,b))`2 IN
        IF rem = zero THEN b
        ELSE Euclidean_gcd_algorithm(R,phi,f_phi)(b,rem)
        ENDIF
  ELSE  Euclidean_gcd_algorithm(R,phi,f_phi)(b,a)
  ENDIF
MEASURE lex2(phi(b), IF a = zero THEN 0 ELSE phi(a) ENDIF)

Euclid_theorem : LEMMA
  FORALL(R:(Euclidean_domain?[T,+,*,zero,one]),
        (phi: [(R - {zero}) -> nat] | Euclidean_pair?(R, phi)),
        (f_phi: [(R),(R - {zero}) -> [(R),(R)]] |
                       Euclidean_f_phi?(R,phi)(f_phi)),
        a: (R), b: (R - {zero}), g : (R - {zero})) :
           gcd?(R)({x : (R) | x = a OR x = b}, g) IFF
           gcd?(R)({x : (R) | x = (f_phi(a,b))`2 OR x = b}, g)

Euclidean_gcd_alg_correctness : THEOREM
  FORALL(R:(Euclidean_domain?[T,+,*,zero,one]),
         (phi: [(R - {zero}) -> nat] | Euclidean_pair?(R, phi)),
         (f_phi: [(R),(R - {zero}) -> [(R),(R)]] |
                        Euclidean_f_phi?(R,phi)(f_phi)),
         a: (R), b: (R - {zero}) ) :
      gcd?(R)({x : (R) | x = a OR x = b},
              Euclidean_gcd_algorithm(R,phi,f_phi)(a,b))
```

Specification 1.8. `gcd` definition for commutative rings - subtheory `ring_gcd_def`

```
gcd?(R)(X: {X | NOT empty?(X) AND subset?(X,R)}, d:(R - {zero})): bool =
    (FORALL a: member(a, X) IMPLIES divides?(R)(d,a)) AND
         (FORALL (c:(R - {zero})):
           (FORALL a: member(a, X) IMPLIES divides?(R)(c,a)) IMPLIES
    divides?(R)(c,d))
```

Now, we show how the correctness of the abstract `Euclidean_gcd_algorithm` is easily inherited, under adequate parameterizations, for the structures of integers $\mathbb{Z}$ and Gaussian integers $\mathbb{Z}[i]$. The lines of reasoning follow those given in discussions on factoration in commutative rings and multiplicative norms in text books (e.g, Section 47 in [10], or Chapter 3, Section 3 in [14]).

The Specification 1.9 presents the case of the Euclidean ring $\mathbb{Z}$. The Euclidean norm $\phi_{\mathbb{Z}}$ is selected as the absolute value while the associated function $f_{\phi_{\mathbb{Z}}}$ is built using the integer division and remainder, specified in the PVS prelude libraries as `div` and `rem`: for $a \in \mathbb{Z}, b \in \mathbb{Z} \setminus \{0\}$, $\texttt{div}(a,b)$ computes the integer division of a by b, and, for $b \in \mathbb{Z}^+ \setminus \{0\}$, $\texttt{rem}(b)(a)$ computes the remainder of a by b. The correctness of the Euclidean algorithm is specified as the corollary `Euclidean_gcd_alg_correctness_in_Z` , that states that for the Euclidean ring of integers $\mathbb{Z}$, and any $i, j \in \mathbb{Z}, j \neq 0$, the parameterized abstract algorithm, `Euclidean_gcd_algorithm[int,+,*,0,1]` satisfies the relation `gcd?[int,+,*,0]`:

$$\texttt{gcd?}[int,+,*,0](\mathbb{Z})(i,j,\texttt{Euclidean\_gcd\_algorithm}[int,+,*,0,1](\mathbb{Z},\phi_{\mathbb{Z}},f_{\phi_{\mathbb{Z}}})(i,j))$$

The formalization of this corollary follows from the theorem of correctness for the abstract Euclidean algorithm, `Euclidean_gcd_alg_correctness` theorem (Specification 1.7), which essentially requires proving that the chosen Euclidean measure $\phi_{\mathbb{Z}}$, and associated function $f_{\phi_{\mathbb{Z}}}$ fulfill the conditions in the definition of Euclidean rings. The latter formalized as lemma `phi_Z_and_f_phi_Z_ok` : $\texttt{Euclidean\_f\_phi?}[\texttt{int},+,*,0](\mathbb{Z},\phi_{\mathbb{Z}})(f_{\phi_{\mathbb{Z}}})$.

Specification 1.9. Correctness of the parameterization of the abstract Euclidean algorithm for the Euclidean ring $\mathbb{Z}$ - subtheory `ring_euclidean_gcd_algorithm_Z`

```
phi_Z(i : int | i /= 0) : posnat =  abs(i)

f_phi_Z(i : int, (j : int | j /= 0)) : [int, below[abs(j)]] =
 ((IF j > 0 THEN ndiv(i,j) ELSE -ndiv(i,-j) ENDIF), rem(abs(j))(i))

phi_Z_and_f_phi_Z_ok  : LEMMA Euclidean_f_phi?[int,+,*,0](Z,phi_Z)(f_phi_Z)

Euclidean_gcd_alg_correctness_in_Z  : COROLLARY
  FORALL(i: int, (j: int | j /= 0)  ) :
    gcd?[int,+,*,0](Z)({x : (Z) | x = i OR x = j},
            Euclidean_gcd_algorithm[int,+,*,0,1](Z, phi_Z,f_phi_Z)(i,j))
```

The Specification 1.10 presents the formalization of correctness of the Euclidean algorithm for the Euclidean ring $\mathbb{Z}[i]$ of Gaussian integers. The Euclidean norm of a Gaussian integer $x = (\texttt{Re}(x) + i\,\texttt{Im}(x)) \in \mathbb{Z}[i]$, $\phi_{\mathbb{Z}[i]}(x)$, is selected as the natural given by the multiplication of x by its conjugate ($\texttt{conjugate}(x) = \texttt{Re}(x) - i\,\texttt{Im}(x)$): $\texttt{Re}(x)^2 + \texttt{Im}(x)^2$. The construction of an adequate associated function $f_{\phi_{\mathbb{Z}[i]}}$ (`f_phi_Zi` in Specification 1.10) requires additional explanations and is specified through the auxiliary function `div_rem_appx`

. For a pair of integers (a, b), $b \neq 0$, this function computes the pair of integers (q, r) such that $a = q\,b + r$, and $|r| \leq |b|/2$; thus, $q\,b$ is the integer closest to a. The equality $a = q\,b + r$ is formalized as lemma div_rev_appx_correctness . Several properties on the field of complex numbers are imported from the PVS complex theory.

Now, we explain the construction of the function $f_{\phi_{\mathbb{Z}[i]}}$. For y, a Gaussian integer and x, a positive integer, let $\texttt{Re}(y) = q_1 x + r_1$ and $\texttt{Im}(y) = q_2 x + r_2$, where (q_1, r_1) and (q_2, r_2) are computed with the auxiliary function div_rem_appx (with respective inputs $(\texttt{Re}(y), x)$ and $(\texttt{Im}(y), x)$). Let $q = q_1 + iq_2$ and $r = r_1 + ir_2$, then $y = qx + r$. Also, notice that if $r \neq 0$ then $\phi_{\mathbb{Z}[i]}(r) \leq \phi_{\mathbb{Z}[i]}(x)$, since $r_1^2 + r_2^2 \leq x^2/2 \leq x^2$. For the case in which x is a non zero Gaussian integer, $\phi_{\mathbb{Z}[i]}(x) > 0$ holds. Then, we can compute $\texttt{div\_rem\_appx}(y\,\texttt{conjugate}(x), x\,\texttt{conjugate}(x))$, obtaining $q, r' \in \mathbb{Z}[i]$ such that $y\,\texttt{conjugate}(x) = q\,(x\,\texttt{conjugate}(x)) + r'$, and $r' = 0$ or $\phi_{\mathbb{Z}[i]}(r') < \phi_{\mathbb{Z}[i]}(x\,\texttt{conjugate}(x))$. Now, select $r = y - q\,x$, then $y = q\,x + r$, and $r\,\texttt{conjugate}(x) = r'$. Finally, when $r \neq 0$, since $\phi_{\mathbb{Z}[i]}(r\,\texttt{conjugate}(x)) < \phi_{\mathbb{Z}[i]}(x\,\texttt{conjugate}(x))$, by application of the lemma phi_Zi_is_multiplicative , we conclude that $\phi_{\mathbb{Z}[i]}(r) < \phi_{\mathbb{Z}[i]}(x)$.

The formalization of correctness of the Euclidean algorithm for Gaussian integers obtained by parameterizations with $\mathbb{Z}[i]$, its Euclidean norm $\phi_{\mathbb{Z}[i]}$ and associated function $f_{\phi_{\mathbb{Z}[i]}}$ follows as the simple corollary Euclidean_gcd_alg_in_Zi in Specification 1.10. This is proved using the correctness of the abstract Euclidean algorithm (Specification 1.7) and lemma phi_Zi_and_f_phi_Zi_ok . The latter states that the Euclidean norm $\phi_{\mathbb{Z}[i]}$ and its associated function $f_{\phi_{\mathbb{Z}[i]}}$ are adequate for the Euclidean ring $\mathbb{Z}[i]$:

$$\texttt{Euclidean\_f\_phi?}[\texttt{complex}, +, *, 0](\mathbb{Z}[i], \phi_{\mathbb{Z}[i]})(f_{\phi_{\mathbb{Z}[i]}})$$

Specification 1.10. Correctness of the parameterization of the abstract Euclidean algorithm for $\mathbb{Z}[i]$ - subtheory ring_euclidean_gcd_algorithm_Zi

```
Zi: set[complex] = {z : complex | EXISTS (a,b:int): a = Re(z) AND b = Im(z)}

Zi_is_ring: LEMMA ring?[complex,+,*,0](Zi)

Zi_is_integral_domain_w_one: LEMMA integral_domain_w_one?[complex,+,*,0,1](Zi)

phi_Zi(x:(Zi) | x /= 0): nat = x * conjugate(x)

phi_Zi_is_multiplicative: LEMMA
   FORALL((x: (Zi) | x /= 0), (y: (Zi) | y /= 0)):
                  phi_Zi(x * y) = phi_Zi(x) * phi_Zi(y)

div_rem_appx(a: int, (b: int | b /= 0)) : [int, int] =
  LET r = rem(abs(b))(a),
      q = IF b > 0 THEN ndiv(a,b) ELSE -ndiv(a,-b) ENDIF  IN
   IF r <= abs(b)/2 THEN (q,r)
   ELSE IF b > 0 THEN (q+1, r - abs(b))
        ELSE (q-1, r - abs(b))
        ENDIF
   ENDIF

div_rev_appx_correctness : LEMMA
```

```
   FORALL (a: int, (b: int | b /= 0)) :
      abs(div_rem_appx(a,b)`2) <= abs(b)/2 AND
      a = b * div_rem_appx(a,b)`1 +  div_rem_appx(a,b)`2

f_phi_Zi(y: (Zi), (x: (Zi) | x /= 0)): [(Zi),(Zi)] =
  LET q = div_rem_appx(Re(y * conjugate(x)), x * conjugate(x))`1 +
          div_rem_appx(Im(y * conjugate(x)), x * conjugate(x))`1 * i,
      r = y - q * x IN (q,r)

 phi_Zi_and_f_phi_Zi_ok: LEMMA
    Euclidean_f_phi?[complex,+,*,0](Zi,phi_Zi)(f_phi_Zi)

 Euclidean_gcd_alg_in_Zi: COROLLARY
  FORALL(x: (Zi), (y: (Zi) | y /= 0)  ) :
      gcd?[complex,+,*,0](Zi)({z :(Zi) | z = x OR z = y},
       Euclidean_gcd_algorithm[complex,+,*,0,1](Zi, phi_Zi,f_phi_Zi)(x,y))
```

4 Related Work

Several formalizations focus on specific ring structures as the ring of integers. Such developments range from simple formalization exercises, such as correctness proofs of gcd algorithms for $\mathbb{Z}$, to elaborated mechanical proofs of the Chinese Remainder theorem for $\mathbb{Z}$. The latter started from Zhang and Hua's RRL (Rewrite Rule Laboratory) mechanization [24], followed by different approaches in Mizar, HOL Light, hol98, and Coq [22], ACL2 [20], and VeriFun [23]. Nevertheless, the general algebraic abstract approach is followed by a few developments. In particular, such an approach is followed in the Isabelle/HOL algebra library (see [1], and [2]); a library that provides a wide range of theorems on mathematical structures, including results on rings, groups, factorization over ideals, rings of integers and polynomial rings. Also, the Lean mathlib library [7] specifies unique factorization domains, prime and irreducible elements in commutative rings, and relations with principal ideal domains. In addition, it specifies the notion of gcd for Euclidean domains and formalizes several properties as the correctness of the extended Euclidean algorithm by applying Bézout's gcd lemma. Nevertheless, mathlib neither includes a general abstract presentation of the Euclidean algorithm nor parameterizations to specific Euclidean domains as given in this paper. A recent extension of mathlib specifies the ring of Witt vectors and formalizes the isomorphism between the ring of Witt vectors over $\mathbb{Z}/p\mathbb{Z}$ and the ring of p-adic integers $\mathbb{Z}p$, for a prime p [6].

In Coq, results about groups, rings, and ordered fields were formalized as part of the FTA project [11]; this work gave rise to the formalization of the Feit and Thompson's proof of the Odd Order Theorem [12]. Also, there are formalizations in Coq of real ordered fields [5], finite fields [19], and rings with explicit divisibility [4]. In Nuprl and Mizar, there are proofs of the Binomial Theorem for rings in [15] and [21], respectively, and a Mizar formalization of the First Isomorphism Theorem for rings [16]. In ACL2, there exists a hierarchy of algebraic structures ranging from setoids to vector spaces that aims the formalization of computer algebra systems [13].

5 Conclusions and Future Work

In contrast to other formalizations that are restricted to specific ring structures, we follow an approach focusing our formalizations on the theory of abstract rings, as done in [7] and [2]. Advantages of following such an approach include increasing the interest of mathematicians in formalizations and having practical general presentations of computational algebraic properties portable to specific ring structures. In particular, in [18], we formalized the Chinese Remainder Theorem for (non-necessarily commutative) rings and obtained as a corollary the CRT version for the ring of integers. This work substantially extended the `algebra` PVS library by specifying Euclidean rings and factorization domains, and formalizing the correspondence between principal ideal domains and unique factorization domains. Also, it proved the correctness of a general Euclidean `gcd` algorithm for Euclidean domains. The usefulness of such an abstraction is made evident through the formalization of simple corollaries stating the correctness of the Euclidean algorithm (parameterized) for the rings of integers and Gaussian integers ($\mathbb{Z}$ and $\mathbb{Z}[i]$).

As future work, we will include the specification of modular arithmetic, and verification of generic versions of Euler's Theorem and Fermat's Little Theorem for Euclidean domains.

References

1. Aransay, J., Ballarin, C., Baillon, M., de Vilhena, P.E., Hohe, S., Kammüller, F., Paulson, L.C.: The Isabelle/HOL Algebra Library. Tech. rep., Isabelle Library, University of Cambridge Computer Laboratory and Technische Universität München (June 2019), https://isabelle.in.tum.de/dist/library/HOL/HOL-Algebra/document.pdf
2. Ballarin, C.: Exploring the structure of an algebra text with locales. Journal of Automated Reasoning **64**, 1093–1121 (2019), https://doi.org/10.1007/s10817-019-09537-9
3. Butler, R., Lester, D.: A PVS *Theory* for Abstract Algebra (2007), http://shemesh.larc.nasa.gov/fm/ftp/larc/PVS-library/pvslib.html, accessed in March 31, 2019
4. Cano, G., Cohen, C., Dénès, M., Mörtberg, A., Siles, V.: Formalized linear algebra over Elementary Divisor Rings in Coq. Logical Methods in Computer Science **12**(2:7), 1–23 (Jun 2016), https://doi.org/10.2168/LMCS-12(2:7)2016
5. Cohen, C., Mahboubi, A.: Formal proofs in real algebraic geometry: from ordered fields to quantifier elimination. Logical Methods in Computer Science **8**(1:2), 1–40 (2012), https://doi.org/10.2168/LMCS-8(1:2)2012
6. Commelin, J., Lewis, R.Y.: Formalizing the ring of Witt vectors. In: 10th ACM SIGPLAN International Conference on Certified Programs and Proofs CPP. pp. 264–277. ACM (2021). https://doi.org/10.1145/3437992.3439919
7. mathlib Community, T.: The Lean Mathematical Library. In: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020. pp. 367–381. ACM (2020), https://doi.org/10.1145/3372885.3373824
8. Dummit, D.S., Foote, R.M.: Abstract Algebra. Wiley, 3 edn. (Jul 2003)

9. Eder, C., Pfister, G., Popescu, A.: On signature-based gröbner bases over euclidean rings. In: Burr, M.A., Yap, C.K., Din, M.S.E. (eds.) Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation, ISSAC 2017, Kaiserslautern, Germany, July 25-28, 2017. pp. 141–148. ACM (2017). https://doi.org/10.1145/3087604.3087614, https://doi.org/10.1145/3087604.3087614
10. Fraleigh, J.B.: A First Course in Abstract Algebra. Pearson, 7th edn. (2003)
11. Geuvers, H., Pollack, R., Wiedijk, F., Zwanenburg, J.: A constructive algebraic hierarchy in Coq. Journal of Symbolic Computation **34**(4), 271–286 (2002), https://doi.org/10.1006/jsco.2002.0552
12. Gonthier, G., Asperti, A., Avigad, J., Bertot, Y., Cohen, C., Garillot, F., Roux, S.L., Mahboubi, A., O'Connor, R., Biha, S.O., Pasca, I., Rideau, L., Solovyev, A., Tassi, E., Théry, L.: A Machine-Checked Proof of the Odd Order Theorem. In: 4th International Conference on Interactive Theorem Proving ITP. Lecture Notes in Computer Science, vol. 7998, pp. 163–179. Springer (2013), https://doi.org/10.1007/978-3-642-39634-2_14
13. Heras, J., Martín-Mateos, F.J., Pascual, V.: Modelling algebraic structures and morphisms in ACL2. Applicable Algebra in Engineering, Communication and Computing **26**(3), 277–303 (Jun 2015), https://doi.org/10.1007/s00200-015-0252-9
14. Hungerford, T.W.: Algebra, Graduate Texts in Mathematics, vol. 73. Springer-Verlag, New York-Berlin (1980), reprint of the 1974 original
15. Jackson, P.B.: Enhancing the Nuprl Proof Development System and Applying it to Computational Abstract Algebra. Ph.D. thesis, Cornell University (1995)
16. Kornilowicz, A., Schwarzweller, C.: The First Isomorphism Theorem and Other Properties of Rings. Formalized Mathematics **22**(4), 291– 301 (2014), https://doi.org/10.2478/forma-2014-0029
17. Lichtblau, D.: Applications of Strong Gröbner Bases over Euclidean Domains. International Journal of Algebra **7**(8), 369–390 (2013)
18. de Lima, T.A., Galdino, A.L., Avelar, A.B., Ayala-Rincón, M.: Formalization of Ring Theory in PVS. J. Autom. Reason. **65**(8), 1231–1263 (2021). https://doi.org/10.1007/s10817-021-09593-0
19. Philipoom, J.: Correct-by-Construction Finite Field Arithmetic in Coq. Master's thesis, Master of Engineering in Computer Science, MIT (2018)
20. Russinoff, D.M.: A Mechanical Proof of the Chinese Remainder Theorem. UTCS Technical Report - no longer available - ACL2 Workshop 2000 TR-00-29, University of Texas at Austin (2000)
21. Schwarzweller, C.: The Binomial Theorem for Algebraic Structures. Journal of Formalized Mathematics **12**(3), 559–564 (2003), http://mizar.org/JFM/Vol12/binom.html
22. Schwarzweller, C.: The Chinese Remainder Theorem, its Proofs and its Generalizations in Mathematical Repositories. Studies in Logic, Grammar and Rhetoric **18**(31), 103–119 (2009), https://philpapers.org/rec/SCHTCR-12
23. Walther, C.: A Machine Assisted Proof of the Chinese Remainder Theorem. Technical Report VFR 18/03, FB Informatik, Technische Universität Darmstadt (2018)
24. Zhang, H., Hua, X.: Proving the Chinese Remainder Theorem by the Cover Set Induction. In: 11th International Conference on Automated Deduction CADE. Lecture Notes in Computer Science, vol. 607, pp. 431–445. Springer (1992), https://doi.org/10.1007/3-540-55602-8_182