A String of Pearls: Proofs of Fermat's Little Theorem

Hing-Lun Chan
joseph.chan@anu.edu.au
Australian National University
and
Michael Norrish
Michael.Norrish@nicta.com.au
Canberra Research Lab., NICTA[1];
also, Australian National University

We discuss mechanised proofs of Fermat's Little Theorem in a variety of styles, focusing in particular on an elegant combinatorial "necklace" proof that has not been mechanised previously. What is elegant in prose turns out to be long-winded mechanically, and so we examine the effect of explicitly appealing to group theory. This has pleasant consequences both for the necklace proof, and also for some of the direct number-theoretic approaches.

1. INTRODUCTION

Fermat's Little Theorem is a famous result in basic number theory. When p is prime, then

$$a^p \equiv a \pmod{p} \qquad \text{for any natural number } a.$$

Though resources like Wikipedia [Wik] provide an extensive range of proofs of this result, it seems that standard practice in interactive proof assistants (*e.g.* Hurd *et al.* [HGF06]) is to use Euler's generalisation, which is number-theoretic. There is good reason for this: the number theory required is actually quite simple, making it easy to establish the result without needing a great deal of background theory. This paper shows, however, how a number of other proofs, some with interesting ideas, can be performed mechanically.

The simplest such proof (the necklace) is based on combinatorics over lists. It is relatively straightforward to mechanise (we use the HOL4 proof assistant [SN08], which has built-in support for lists), but aspects of the proof become smoother when it is rephrased in the language of group theory. This required background does not represent a particularly onerous burden for mechanisation. Indeed, the more group theory one has to hand, the more polished the proof becomes.

We also examine the effect of using group theory in number-theoretic approaches.

Overview. The rest of the paper is structured as follows. Section 2 is devoted to the history of Fermat's Little Theorem. In Sections 3 and 4, we describe both the standard number-theoretic proof, and Golomb's combinatorial necklace proof [Gol56], and their mechanisation. In Section 5, we discuss how the required

[1]NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

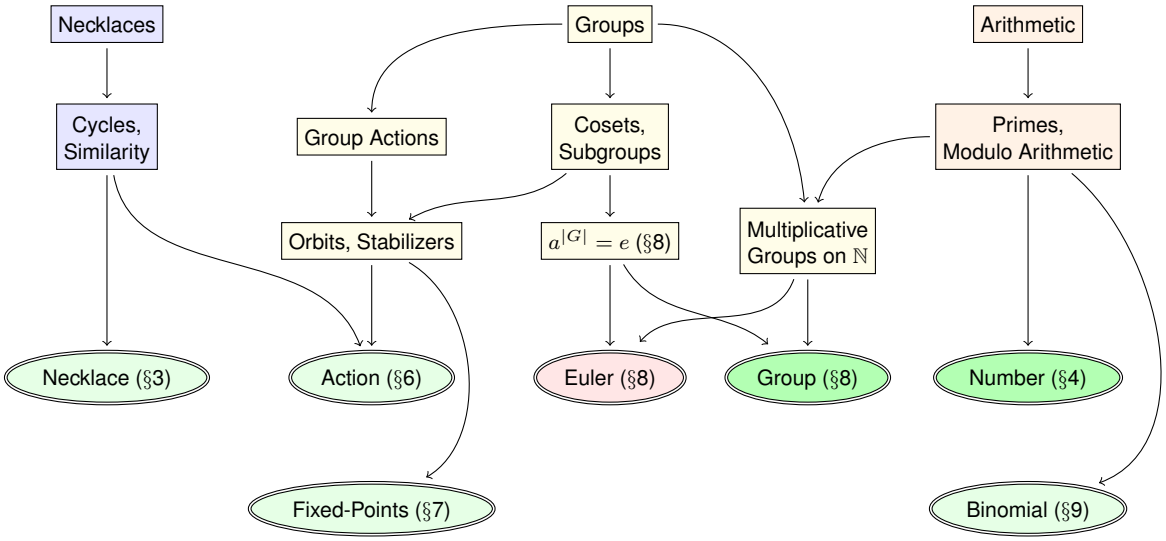

Fig. 1. Theory dependencies for proofs of Fermat's Little Theorem (FLT). Double-lined ellipses indicate proofs of FLT discussed in the corresponding section of the paper. The leftmost Necklace, rightmost Number and Binomial are direct proofs; others use group theory. Binomial and combinatoric results (in light green) are of the form $a^p \equiv a \pmod p$, which is equivalent (when $0 < a < p$ for prime p, see Theorem 9 and its footnote) to number-theoretic results (in dark green) of the form $a^{p-1} \equiv 1 \pmod p$. Euler (in light red) is of the form $a^{\varphi(n)} \equiv 1 \pmod n$.

(rather basic) group theory is mechanised, before showing how this theory can be applied to the necklace proof (in Section 6 and Section 7), and to the number-theoretic proof (in Section 8). For the sake of completeness, we deliver an induction proof in Section 9, and we conclude in Section 10, including a comparison of the different approaches in terms of their complexity. One of the HOL4 proofs of Fermat's Little Theorem is included in Appendix.

Figure 1 gives a graphical summary of the logical dependencies underlying all of the proofs we discuss. The graph falls into three parts:

— The leftmost path shows Golomb's necklace proof, a combinatorial proof based on rotations of lists.

— The rightmost paths show two elementary number-theoretic proofs commonly found in textbooks.

— The central paths are of various group-theoretic proofs. The first path from the left is the application of group theory (*via* the notions of action, orbit, stabilizer, and fixed-points) to the necklace proof. The others show the proof of the group-theoretic analogue of Fermat's Little Theorem, followed by the derivation of specific results in the domain of natural numbers.

HOL4 Notation and Theorems. All statements starting with a turnstile ($\vdash$) are HOL4 theorems, automatically pretty-printed to LaTeX from the relevant theory in the HOL4 development. Notation specific to this paper is explained as it is introduced. Otherwise, HOL4 supports a notation that is a generally pleasant combination of quantifiers ($\forall$, $\exists$) and functional programming (λ for function abstraction, and juxtaposition for function application).

Lists are written between square brackets, *e.g.*, `[1; 2; 3]`. The length of a list ℓ is written $|\ell|$. The concatenation of lists ℓ_1 and ℓ_2 is written ℓ_1 `++` ℓ_2. Applying a function f over an index list $[0; \ldots ;(n-1)]$ of n terms generates `GENLIST` f n, and `SUM` ℓ gives the total sum of all terms in list ℓ.

Sets are written between braces, also allowing comprehensions such as $\{x \mid x < 6\}$. Sets support standard operations such as cardinality (also written with vertical bars: $|\{3; 5\}| = 2$), union (∪), intersection (∩), and difference (\). We write IMAGE f s for the image of set s under function f, and BIJ f s_1 s_2 means that function f is a bijection between sets s_1 and s_2. The term R equiv_on s means that R is an equivalence relation on the set s, and partition R s denotes the set of subsets of s that are partitions with respect to an equivalence relation R.

The use of overloading in HOL4's pretty-printing can conceal "implicit" parameters such as the underlying group in a term such as $x \times y$. This latter may denote multiplication of natural numbers x and y, or denote the result of the group operation applied to elements x and y of a group g. We hope the context makes it clear which operation is being used.

Our Contribution. As already noted, Fermat's Little Theorem has been mechanised a number of times before, *e.g.*, in Coq [Oos], ACL2 [Rus07], Matita [AA08] and HOL Light [Har11]. The minimal group theory we used and mechanised is also very standard.[1] Our contribution (first appearing in the earlier [CN12]) is the mechanisation of the necklace proof, in direct and group-theoretic styles (we believe both to be entirely novel). We also compare these proofs with the standard number-theoretic approaches.

Availability. HOL4 proof scripts can be found at `http://bitbucket.org/jhlchan/hol/src`. The linearised scripts (discussed in Section 10) are those beginning with prefix `All` in the `fermat` directory. Proofs as they were developed (in various separate theories) are laid out in sub-directories below `fermat`.

2. HISTORY OF FERMAT'S LITTLE THEOREM

Fermat stated this result in a letter to Frénicle de Bessy dated October 18, 1640 [Mah94, Chapter VI]:

> *Without exception, every prime number measures one of the powers reduced by unity of any progression whatever, and the exponent of the said power is a sub-multiple of the given prime number reduced by unity. [. . .] This proposition is generally true for all progressions and for all prime numbers; the proof of which I would send to you, if I were not afraid to be too long.*

In modern notation, this is:

$$a^{p-1} \equiv 1 \mod p \quad \text{for prime } p \text{ and } a \text{ coprime to } p$$

which is equivalent to the usual statement (see footnote of Theorem 9). A proof based on mathematical induction via the binomial theorem (later Section 9) was given by Leibniz (1646–1716), a contemporary of Fermat's. However, this was in an unpublished and undated manuscript discovered only in 1894 among the archives in Hannover Library [MOS06, Chapter 4], almost two centuries later. Leibniz wrote in this manuscript that he knew of the proof before 1683.

In 1729, Goldbach asked Euler if he could verify one of Fermat's many claims. This led Euler to study Fermat's work, especially in number theory. In 1736, Euler published a proof of Fermat's Little Theorem, which was essentially Leibniz's proof using the binomial theorem. His second proof, published in 1747, was a variation on the same theme, still based on induction. However, in 1758, Euler published a third proof, which avoided using the binomial theorem, and employed essentially the modern group-theoretic viewpoint. Based on his third proof, Euler published a generalisation (later Section 8.1) of Fermat's result in 1760 [San03].

Although Fermat did not leave behind any written evidence of his proof, André Weil's attempted reconstruction [Wei84, Chapter II, $IV] suggested that Fermat may have started with a proof similar to Leibniz, but then later conceived of the same group-theoretic idea as Euler. For another reconstruction of Fermat's proof, see Burn [Bur02].

[1] The Orbit-Stabiliser theorem has not been mechanised before in HOL, but this is a minor contribution given the existing work in other systems such as Coq.

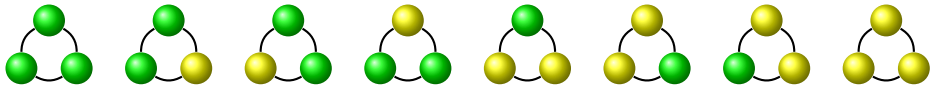

Fig. 2. Necklaces with 3 beads in 2 colours. The first and last are *monocoloured* necklaces. The other *multicoloured* necklaces are divided into 2 parts: those with one yellow bead and those with two yellow beads. Multicoloured necklaces in each part can cycle only among themselves. Note that, for 3 beads, each part consists of 3 necklaces. Hence the number of multi-2-coloured necklaces with 3 beads (which is $2^3 - 2 = 8 - 2 = 6$) is divisible by 3.

3. THE NECKLACE PROOF

Due to the intrinsic simplicity of Fermat's Little Theorem, others devised better and shorter proofs of this basic result. Dickson, in his authoritative treatise *History of the Theory of Numbers* [Dic19, Chapter 3], thoroughly documented all known proofs of this Fermat's result up to 1919. Among them was this nice combinatorial proof by Julius Petersen from 1872:

> *Take p elements from q with repetitions in all ways, that is, in q^p ways. The q sets with elements all alike are not changed by a cyclic permutation of the elements, while the remaining $q^p - q$ sets are permuted in sets of p [when p is prime]. Hence p divides $q^p - q$.*

This idea is the basis of the *Necklace proof* of S. W. Golomb [Gol56], which has since been rediscovered or discussed by many others (*e.g.*, Smyth [Smy86], Rouse [Rou03], Anderson [ABR05], Evans [HE], and Conrad [Con08]).

The necklaces of Golomb's proof are the p elements drawn from a set of cardinality q. Where Peterson has cyclic permutations, Golomb's version adds the image of rotating beads on a necklace (Figure 2).

3.1 Necklaces and Colours

Consider the set N of necklaces[2] of length n (*i.e.*, n beads) with a colours (*i.e.*, a choices for a bead's colour). Since a bead can have any of the a colours, and there are n beads in total, the total number of necklaces is $|N| = a^n$. Of these necklaces, the *monocoloured* necklaces are those with the same colour for all beads; the others are *multicoloured* necklaces.

Let S (for single) denote the set of monocoloured necklaces, and M (for multiple) denote the multicoloured necklaces. Clearly, $N = S \cup M$, and $S \cap M = \emptyset$ — that is, S and M are disjoint, and they form a partition of the set of necklaces N. Since there is only 1 monocoloured necklace for each colour, the number of monocoloured necklaces $|S|$ is just a. Given that the two types of necklaces partition the whole set, the number of multicoloured necklaces $|M|$ is equal to $|N| - |S| = a^n - a$.

HOL Implementation. Let $(\texttt{necklace}\ n\ a)$ be the set of necklaces of length n with a colours. The HOL definition is

$$\vdash \texttt{necklace}\ n\ a = \{\ell \mid |\ell| = n \wedge \texttt{set}\ \ell \subseteq \texttt{count}\ a\}$$

Our necklaces' beads are just natural numbers, and the definition requires that the "colours" of the necklace are simply drawn from the set $(\texttt{count}\ a)$: the set of natural numbers less than a.

Simple properties of the set $(\texttt{necklace}\ n\ a)$ readily follow from the definition:

$$\vdash \texttt{FINITE (necklace}\ n\ a\texttt{)}$$
$$\vdash |\texttt{necklace}\ n\ a| = a^n$$

The monocoloured and multicoloured necklaces are defined thus:

$$\vdash \texttt{monocoloured}\ n\ a = \{\ell \mid \ell \in \texttt{necklace}\ n\ a \wedge (\ell \neq \texttt{[]} \Rightarrow \texttt{SING (set}\ \ell\texttt{))}\}$$
$$\vdash \texttt{multicoloured}\ n\ a = \texttt{necklace}\ n\ a \setminus \texttt{monocoloured}\ n\ a$$

[2] For ease of counting, necklaces here are fixed in space. The rotation equivalence of necklaces will be treated in Section 3.2.

where `SING (set ℓ)` means that the set of list elements `(set ℓ)` is a singleton. The cardinality results for these sets are straightforward:

$\vdash$ `0 < n ⇒ |monocoloured n a| = a`
$\vdash$ `0 < n ⇒ |multicoloured n a| =` $a^n - a$

In order to show that the last expression $a^n - a$ is divisible by n when length n is prime, we need to know something more about the multicoloured necklaces, especially how an equivalence relation involving cyclic permutations partitions the set.

3.2 Cycles

Necklaces are represented by lists of length n. Following the imagery, it is natural to think of them as being joined from end to end. We define a `cycle` operation on lists:

$\vdash$ `cycle n ℓ = FUNPOW (λℓ. DROP 1 ℓ ++ TAKE 1 ℓ) n ℓ`

The expression `DROP n ℓ` discards the first n elements of list ℓ, returning whatever remains, while `TAKE n ℓ` returns the first n elements (for the empty list `[]`, `TAKE` and `DROP` both return `[]`). By putting $n = 1$, this is chopping off the first bead, shifting it to the other end and adding it back. Therefore `DROP 1 ℓ ++ TAKE 1 ℓ` represents a rotation by 1 bead position. Then `FUNPOW` just repeats this operation n times.[3] These elementary facts about `cycle` follow immediately:

$\vdash$ `cycle 0 ℓ = ℓ` (CYCLE_0)
$\vdash$ `cycle n (cycle m ℓ) = cycle (n + m) ℓ` (CYCLE_ADD)

Applying `cycle` on a necklace results in another necklace, of the same length and colours:

$\vdash$ `ℓ ∈ necklace n a ⇒ ∀k. cycle k ℓ ∈ necklace n a`
$\vdash$ `|cycle n ℓ| = |ℓ|`
$\vdash$ `set (cycle n ℓ) = set ℓ`

As a result, `cycle` of a monocoloured necklace is still monocoloured, and `cycle` of a multicoloured necklace is still multicoloured, as expected.

We can reason about cycles with modular arithmetic:

$\vdash$ `ℓ ≠ [] ⇒ cycle n ℓ = cycle (n MOD |ℓ|) ℓ` (CYCLE_MOD_LENGTH)
$\vdash$ `ℓ ≠ [] ⇒ cycle m (cycle n ℓ) = cycle ((m + n) MOD |ℓ|) ℓ`

And ultimately, a cycle can come full-circle, in multiples, or can be undone by another cycle:

$\vdash$ `cycle |ℓ| ℓ = ℓ` (CYCLE_BACK)
$\vdash$ `cycle n ℓ = ℓ ⇒ ∀m. cycle (m × n) ℓ = ℓ` (CYCLE_MULTIPLE_BACK)
$\vdash$ `n ≤ |ℓ| ⇒ cycle (|ℓ|−n) (cycle n ℓ) = ℓ` (CYCLE_INV)

Already, one can see the possible connections to group theory.

3.3 Similarity and Partitions

We shall say two necklaces ℓ_1, ℓ_2 are *similar*, denoted ℓ_1 `==` ℓ_2, when:

$\vdash$ ℓ_1 `==` ℓ_2 $\iff$ `∃n. ℓ2 = cycle n ℓ1`

That is, ℓ_1 can cycle to ℓ_2 because they consist of the same beads in cyclic order.
The following properties of `(==)` follow from properties of `cycle`:

[3] This definition of `cycle n ℓ` using `FUNPOW` makes sense for all n, whereas a definition using `TAKE n` and `DROP n` would only work when $n \leq |\ell|$.

$\vdash \ell == [] \vee [] == \ell \iff \ell = []$
$\vdash \ell_1 == \ell_2 \Rightarrow |\ell_1| = |\ell_2|$

With a little more effort, the fact that (==) is an equivalence relation can be proved:

$\vdash \ell == \ell$
$\vdash \ell_1 == \ell_2 \Rightarrow \ell_2 == \ell_1$
$\vdash \ell_1 == \ell_2 \wedge \ell_2 == \ell_3 \Rightarrow \ell_1 == \ell_3$

The key for reflexivity is CYCLE_0, for symmetry is CYCLE_INV, for transitivity is CYCLE_ADD. Let us denote the equivalence classes under (==) by associates:

$\vdash$ associates $x = \{y \mid x == y\}$

As (==) is an equivalence relation, associates partition the set of necklaces. This partitioning has a particularly simple structure when the necklace length n is prime.

3.4 Multicoloured Necklaces with Prime Length

First, an important result about values that "cycle back" and their greatest common divisor:

THEOREM 1. *If two values m, n can cycle back, the value* gcd m n *can also cycle back.*

$\vdash$ cycle m ℓ = ℓ $\wedge$ cycle n ℓ = ℓ $\Rightarrow$ cycle (gcd m n) ℓ = ℓ

PROOF. If n = 0, then cycle (gcd m 0) ℓ = cycle m ℓ = ℓ by assumption. Otherwise, we can use Bézout's identity, called LINEAR_GCD in HOL library, which states that if $n \neq 0$, then there exist p and q such that $p \times n = q \times m +$ gcd m n, and reason:

```
    cycle (gcd m n) ℓ
  = cycle (gcd m n) (cycle (q × m) ℓ)      by CYCLE_MULTIPLE_BACK
  = cycle (gcd m n + q × m) ℓ              by CYCLE_ADD
  = cycle (p × n) ℓ                        by LINEAR_GCD
  = ℓ                                      by CYCLE_MULTIPLE_BACK
```

□

A distinguishing feature of monocoloured necklaces is:

THEOREM 2. *A necklace ℓ is monocoloured iff* cycle 1 ℓ = ℓ.

$\vdash 0 < n \wedge 0 < a \wedge \ell \in$ necklace n $a \Rightarrow$
$(\ell \in$ monocoloured n $a \iff$ cycle 1 ℓ = $\ell)$

PROOF. Since a monocoloured necklace ℓ has all its beads the same colour, shifting 1 bead makes no difference, hence cycle 1 ℓ = ℓ. Conversely, given a necklace ℓ with cycle 1 ℓ = ℓ, applying CYCLE_MULTIPLE_BACK, we have ℓ = cycle 2 ℓ = cycle 3 ℓ = ... As lists, head of ℓ is the first bead, head of cycle 1 ℓ is the second bead, head of cycle 2 ℓ is the third bead, *etc*. Since these cycle lists are all the same, and equal lists mean equal heads, all beads have the same colour, making the necklace ℓ monocoloured. □

We proceed to find the size of associates of multicoloured necklaces with prime length:

THEOREM 3. *For multicoloured necklaces ℓ with prime $|\ell|$ = p, the* cycle *map from* count p *to* associates ℓ *is injective.*

$\vdash$ prime p $\wedge$ $\ell \in$ multicoloured p $a \Rightarrow$
INJ (λn. cycle n ℓ) (count p) (associates ℓ)

PROOF. This is to show that, for all $x < p$ and $y < p$, `cycle` x ℓ = `cycle` y $\ell \Rightarrow x = y$. Suppose this is not the case. Then there are $x \neq y$ such that there is a common multicoloured necklace ℓ' = `cycle` x ℓ = `cycle` y ℓ. Note that both necklaces ℓ' and ℓ are multicoloured with same length p (Section 3.2). Without loss of generality, assume $x < y$. Then $y = d + x$, where difference $d > 0$ and $d < p$ (since both $x < p$ and $y < p$). Hence `cycle` d ℓ' = `cycle` d (`cycle` x ℓ) = `cycle` (d + x) ℓ = `cycle` y ℓ = ℓ'. With `cycle` d ℓ' = ℓ', and from `CYCLE_BACK` we have `cycle` p ℓ' = ℓ', hence `cycle` (`gcd` d p) ℓ' = ℓ' by Theorem 1. But `gcd` d $p = 1$ for prime $p, 0 < d < p$. This implies the multicoloured necklace ℓ' has `cycle` 1 ℓ' = ℓ', which is a contradiction in view of Theorem 2. □

THEOREM 4. *For multicoloured necklaces ℓ with $|\ell|$ = n (prime or non-prime), the* `cycle` *map from* `count` n *to* `associates` ℓ *is surjective.*

$\vdash$ $\ell \in$ `multicoloured` n $a \Rightarrow$ `SURJ` (λk. `cycle` k ℓ) (`count` n) (`associates` ℓ)

PROOF. This is because, if a necklace ℓ' is similar to ℓ, there is a k such that ℓ' = `cycle` k ℓ. By `CYCLE_MOD_LENGTH`, ℓ' = `cycle` (k `mod` n) ℓ, and so k `mod` n is in the range of `count` n. □

THEOREM 5. *For multicoloured necklaces ℓ with prime $|\ell|$ = p, their associates have size p.*

$\vdash$ `prime` $p \wedge \ell \in$ `multicoloured` p $a \Rightarrow$ |`associates` ℓ| = p

PROOF. Since the `cycle` map is surjective in general (Theorem 4), and injective when the necklace length is prime (Theorem 3), there is a bijection between `count` p and `associates` ℓ for multicoloured necklaces ℓ when $|\ell|$ = p is prime. The result follows from this bijection between finite sets. □

This leads directly to the following mechanisation of the necklace proof of

THEOREM 6. *Fermat's Little Theorem.*

$\vdash$ `prime` $p \Rightarrow p$ `divides` $a^p - a$

PROOF. For prime p, the multicoloured necklaces $\ell \in$ `multicoloured` p a are *"permuted in sets of p"*, as claimed by Julius Petersen (Section 3), since |`associates` ℓ| = p by Theorem 5. Recall that `associates` ℓ are the equivalence classes of (==) on `multicoloured` p a (Section 3.3). Since equivalence classes form a partition, and here they all have the same size p, we have:

$$|\texttt{multicoloured}\ p\ a| = p \times |\texttt{partition}\ (==)\ (\texttt{multicoloured}\ p\ a)|$$

As p is prime, $0 < p$, and |`multicoloured` p a| = $a^p - a$ (Section 3.1). Combining these results we have p `divides` $a^p - a$ by definition of `divides`. □

4. DIRECT NUMBER-THEORETIC PROOF

The proof of Fermat's Little Theorem given in most textbooks, and also that given in various theorem-proving systems, is number-theoretic, based on properties of modulo arithmetic. In particular, modulo prime multiplication has some special properties. The first one, usually referred to as Euclid's Lemma, is that a prime divides a product iff the prime divides a factor. In terms of modulo arithmetic, this is:

$\vdash$ `prime` $p \Rightarrow (x \times y \equiv 0 \pmod{p} \iff x \equiv 0 \pmod{p} \vee y \equiv 0 \pmod{p})$

This gives a very useful trick for modulo computation:

THEOREM 7. *Left-cancellation of a non-zero factor is possible in prime modulo arithmetic.*

$\vdash$ `prime` $p \wedge x \times y \equiv x \times z \pmod{p} \wedge x \not\equiv 0 \pmod{p} \Rightarrow y \equiv z \pmod{p}$

PROOF.

$$\begin{array}{ll} & x \times y \equiv x \times z \pmod{p} \\ \Rightarrow & x \times (y-z) \equiv 0 \pmod{p} \qquad \text{by distribution law} \\ \Rightarrow & y-z \equiv 0 \pmod{p} \qquad \text{by Euclid's Lemma, } x \not\equiv 0 \pmod{p} \\ \Rightarrow & y \equiv z \pmod{p} \end{array}$$

□

DEFINITION 1. *Let the* residues *of prime p be the non-zero numbers less than p:* $\{1 \ldots p-1\}$.

Take any a from the residues of p, and consider the various values of $a \times x \mod p$, for all x also a residue of p. In HOL, this is denoted by a *row* operation:

DEFINITION 2. $\vdash$ `row p a x = a×x mod p`

THEOREM 8. *The row products form a permutation of the residues for prime modulo.*

$\vdash$ `prime p ∧ a ∈ {1...p−1} ⇒ {1...p−1} = IMAGE (row p a) {1...p−1}`

PROOF. The `IMAGE` on the right-hand side is equivalent to $\{a \times x \mod p \mid 1 \leq x \wedge x < p\}$. The possible remainders under modulo p are $0, 1, \ldots p-1$. Since a prime p has no proper factors, and both a and x are less than p, the product $a \times x$ cannot be the prime p, nor any multiple of the prime p. Hence the remainder, $a \times x \mod p$ cannot be zero, making this result one of the residues of p. The possible values are distinct because if $a \times x \equiv a \times y \pmod{p}$, then $x \equiv y \pmod{p}$ by left-cancellation of non-zero a. So the right-hand side, the row products, is just a permutation of the left-hand side, the residues of p. □

This is the key for the number-theoretic proof of

THEOREM 9. *Fermat's Little Theorem (equivalent form[4])*

$\vdash$ `prime p ∧ a ∈ {1...p−1} ⇒` $a^{p-1} \equiv 1 \pmod{p}$

PROOF. Consider multiplying all numbers (denoted by the symbol $\prod$) from each of these finite sets:
(1) the residues $\{1 \ldots p-1\}$ and
(2) its row products `IMAGE (row p a) {1...p−1}`.

Clearly, the first one is a factorial:

$\vdash \prod$ `{1...p−1} = (p−1)!`

For the second one, since for numbers the order of multiplication does not affect the product, all the factors a of (`row p a`) (Definition 2) can be collected together, so we have:

$\vdash$ `prime p ∧ a ∈ {1...p−1} ⇒`
$\prod$ `(IMAGE (row p a) {1...p−1})` $\equiv a^{p-1} \times$ `(p−1)!` $\pmod{p}$

As the underlying sets are the same due to permutation (Theorem 8), the two products under modulo p are identical:

$\vdash$ `prime p ∧ a ∈ {1...p−1} ⇒ (p−1)!` $\equiv a^{p-1} \times$ `(p−1)!` $\pmod{p}$

A prime p has no proper factor, and `(p−1)!` has, as factors, of all the numbers less than p, so `(p−1)!` $\not\equiv 0 \pmod{p}$. Applying non-zero left-cancellation of modulo p multiplication gives Fermat's Little Theorem. □

[4]To show $a^p \equiv a \pmod{p}$ for all a, it is sufficient to show this congruence for $0 \leq a < p$, the possible remainders under modulo p. If $a = 0$ this is trivial. If $a \neq 0$, `gcd a p = 1` since p is prime and $0 < a < p$. This allows left-cancellation of non-zero a on both sides (later Theorem 7), giving the equivalent form $a^{p-1} \equiv 1 \pmod{p}$.

5. GROUP THEORY

The combinatorial necklace proof and the number-theoretic proof may appear unrelated, but there is an underlying algebra behind both proofs, that of group theory. The algebra gives us:

— the cycles and similarities in the necklace proof;

— the factor cancellation in the number-theoretic proof; and

— an insight, allowing a modest generalisation.

The discussion that follows is an expansion of this theme.

We mechanise the necessary theorems from group theory following an existing mechanisation in the HOL distribution by Joe Hurd [HGF06].[5] We have a predicate `Group` g on a record of four fields. Each field can be selected with "dot notation" so that the group operation is g.`mult`, the inverse is g.`inv`, the identity is g.`id`, and the carrier set is g.`carrier`. We abuse notation so that the group g's operation applied to x and y argument appears as $x \times y$, the identity appears as e, the inverse applied to an argument x appears as x^{-1}, and also `G` and `H` stand for the carrier sets of groups g and h respectively:

$$\begin{aligned}&\vdash \texttt{Group}\ g \iff \\ &\quad e \in \texttt{G} \wedge (\forall x\ y\texttt{::(G)}.\ x \times y \in \texttt{G}) \wedge (\forall x\texttt{::(G)}.\ x^{-1} \in \texttt{G}) \wedge \\ &\quad (\forall x\texttt{::(G)}.\ e \times x = x) \wedge (\forall x\texttt{::(G)}.\ x^{-1} \times x = e) \wedge \\ &\quad \forall x\ y\ z\texttt{::(G)}.\ x \times y \times z = x \times (y \times z)\end{aligned}$$

The double-colon notation (*e.g.*, $\forall x\ y\texttt{::(G)}.\ P\ x\ y$) denotes a restriction on all the preceding bound variables (x and y here) requiring them to be in the set `G`.

In this mechanisation, typical results appear with the `Group` predicate as a side-condition.

$$\vdash \texttt{Group}\ g \Rightarrow \forall x\ y\ z\texttt{::(G)}.\ x \times y = x \times z \iff y = z$$
$$\vdash \texttt{Group}\ g \Rightarrow \forall x\ y\ z\texttt{::(G)}.\ x \times y = z \iff x = z \times y^{-1}$$
$$\vdash \texttt{Group}\ g \Rightarrow \forall x\texttt{::(G)}.\ (x^{-1})^{-1} = x$$

This is perhaps not the slickest possible presentation of abstract algebra, even within the constraints of HOL4's simple type theory, but it is both well-understood and sufficient for our purposes.

Group exponentiation is defined *via* primitive recursion, giving us the usual properties:

$$\vdash \texttt{Group}\ g \wedge x \in \texttt{G} \Rightarrow x^0 = e$$
$$\vdash \texttt{Group}\ g \wedge x \in \texttt{G} \Rightarrow x^1 = x$$
$$\vdash \texttt{Group}\ g \wedge x \in \texttt{G} \Rightarrow x^{m \times n} = (x^m)^n$$
$$\vdash \texttt{Group}\ g \wedge x \in \texttt{G} \Rightarrow (x^n)^{-1} = (x^{-1})^n$$

We write $h \leq g$ to mean that h is a subgroup of g, and define the *coset* of a set X with respect to a group element a (normally written aX) to be

$$\vdash \texttt{coset}\ g\ X\ a = \texttt{IMAGE}\ (\lambda z.\ a \times z)\ X$$

The cosets of a subgroup's carrier are important because of these standard results:

THEOREM 10. *Subgroup cosets partition the group's carrier set, by the following equivalence relation:*

$$\vdash \texttt{Group}\ g \wedge h \leq g \Rightarrow \texttt{coset}\ g\ \texttt{H}\ \texttt{equiv\_on}\ \texttt{G}$$

THEOREM 11. *Each coset of a subgroup is in bijection with the subgroup itself.*

$$\vdash \texttt{Group}\ g \wedge h \leq g \wedge a \in \texttt{G} \Rightarrow \texttt{BIJ}\ (\lambda x.\ a \times x)\ \texttt{H}\ (\texttt{coset}\ g\ \texttt{H}\ a)$$

[5] The source code of a prior HOL mechanisation of group theory by Elsa L. Gunter [Gun89] is not generally available.

This bijection allows determination of the size of subgroup cosets:

THEOREM 12. *For a finite subgroup, the size of its coset equals the size of subgroup itself:*

$\vdash$ `Group` $g \wedge h \leq g \wedge a \in$ `G` $\wedge$ `FINITE H` $\Rightarrow$ $|$`coset` g `H` $a|$ = $|$`H`$|$

Therefore the subgroup cosets partition consists of equal-size chunks, leading to Lagrange's Identity:

$\vdash$ `FiniteGroup` $g \wedge h \leq g \Rightarrow$ $|$`G`$|$ = $|$`H`$| \times |$`partition (coset` g `H) G`$|$

and Lagrange's Theorem on cardinality of subgroups:

$\vdash$ `FiniteGroup` $g \wedge h \leq g \Rightarrow$ $|$`H`$|$ `divides` $|$`G`$|$

6. GROUP THEORY APPLIED TO THE NECKLACE PROOF

The group-theoretic version of the necklace proof requires a little more theory than the basic development of the preceding section. We shall use the group $\mathbb{Z}_n^+$, which is the additive group over the natural numbers less than n. This group's binary operation is addition modulo n, and its identity is zero.

6.1 Group Actions

DEFINITION 3. *Let g be a group over a set of elements of type α, X be a set of elements of type β, and (infix) $\circ$ a function of type α -> β -> β. The mapping $(\circ)$ is called a* group action *from g to X, (written* `action (∘) g X`*), if these three conditions are satisfied:*

— *Closure:* $a \in G \wedge x \in X \Rightarrow a \circ x \in X$

— *Identity:* $x \in X \Rightarrow e \circ x = x$

— *Composition:* $a, b \in G \wedge x \in X \Rightarrow a \circ (b \circ x) = (a \times b) \circ x$

The HOL definition is

```
⊢ action (∘) g X ⟺
    ∀x.
      x ∈ X ⇒
        (∀a::(G). a ∘ x ∈ X) ∧ e ∘ x = x ∧
        ∀a b::(G). a ∘ b ∘ x = (a × b) ∘ x
```

The set X above is called the *target*. We can picture a target point $x \in X$ being *acted upon* by the group elements. Alternatively, we say that point x can *reach* another point $a \circ x$ for some $a \in G$. If $a \circ x = x$, we say that the group element a leaves the point x *fixed*. This leads to the following:

DEFINITION 4. *For $x \in X$, the set of target points it can reach form its* **`orbit`**.

DEFINITION 5. *For $x \in X$, the set of group elements that leave it fixed form its* **`stabilizer`**.

For example, $\mathbb{Z}_n^+$ acts on the set of necklaces of length n, with `cycle` being an action from $\mathbb{Z}_n^+$ to the necklaces. Each monocoloured necklace always cycles to itself. Thus its orbit consists of itself only, and its stabilizer is all of the group's carrier. For each multicoloured necklace, cycling brings it to another (similar) multicoloured necklace. Since $|\mathbb{Z}_n^+| = n$, its orbit contains at most n reachable points in the target. Generally, more reachable points give a larger orbit, and the corresponding stabilizer is smaller. In the extreme case when the orbit has n distinct target points, the stabilizer contains just the group identity.

This is a hint that sizes of orbits and stabilizers may have a relationship — an idea we shall explore.

HOL Implementation. The HOL definitions of these concepts pick up multiple parameters, so that, for example, the `reach` relation is not simply a binary notion but has to include explicit parameters for the action and the governing group. Similar extra parameters are required for `orbit` and `stabilizer` definitions:

$\vdash$ `reach (∘) g x y` $\Longleftrightarrow$ $\exists a.$ $a \in$ `G` $\wedge$ $a \circ x = y$
$\vdash$ `orbit (∘) g X x` $= \{y \mid y \in X \wedge$ `reach (∘) g x y`$\}$
$\vdash$ `stabilizer (∘) g x` $= \{a \mid a \in$ `G` $\wedge$ $a \circ x = x\}$

For presentational reasons, we shall assume fixed operation $(\circ)$, group g and target set X in much of what follows, and use the following abbreviations in prose and HOL theorems:

— *orbit* x for `orbit (∘) g X` x, and
— *stabilizer* x for `stabilizer (∘) g` x.

6.2 Action Basics

Properties of group actions blend nicely with properties of groups, as shown by these basic results.

THEOREM 13. *Reachability is an equivalence relation on the target set.*

$\vdash$ `Group` $g \wedge$ `action (∘)` g $X \Rightarrow$ `reach (∘)` g `equiv_on` X

PROOF. Let $x \sim y$ stand for `reach (∘)` g x y. By action identity: $e \circ x = x$, hence $x \sim x$, or `reach` is *reflexive*. If $a \in G$ moves point x to y: $a \circ x = y$, then $a^{-1} \in G$ moves y to x: $a^{-1} \circ y = a^{-1} \circ (a \circ x) = (a^{-1} \times a) \circ x = e \circ x = x$, hence $x \sim y \Rightarrow y \sim x$, or `reach` is *symmetric*. If $a \circ x = y$, and $b \circ y = z$, then by action composition: $(b \times a) \circ x = b \circ (a \circ x) = b \circ y = z$, hence $x \sim y \wedge y \sim z \Rightarrow x \sim z$, or `reach` is *transitive*. Thus `reach` is an equivalence relation. □

The *orbits* are equivalence classes of `reach`, and they form a partition of the target set X. This provides another characterisation of `orbit` using the action mapping $(\circ)$:

$\vdash$ `Group` $g \wedge$ `action (∘)` g $X \wedge x \in X \Rightarrow$ *orbit* $x = \{a \circ x \mid a \in$ `G`$\}$

This characterisation is the basis for the visualization of orbits later.

For a point $x \in X$, *stabilizer* x is not only a subset of G, the group carrier, but also a subgroup of group g, denoted by `StabilizerGroup (∘)` g x:

THEOREM 14. *The stabilizer of a point in the target set forms a subgroup.*

$\vdash$ `action (∘)` g $X \wedge x \in X \wedge$ `Group` $g \Rightarrow$ `StabilizerGroup (∘)` g $x \leq g$

PROOF. If two elements $a, b \in G$ both fix a point $x \in X$, *i.e.*, $a \circ x = x$ and $b \circ x = x$, then by action composition: $(a \times b) \circ x = a \circ (b \circ x) = a \circ x = x$. Therefore, the stabilizer is a closed subset of G. The identity e is in the stabilizer by action identity: $e \circ x = x$. If a is in the stabilizer, its inverse a^{-1} is also in the stabilizer: $a^{-1} \circ x = a^{-1} \circ (a \circ x) = (a^{-1} \times a) \circ x = e \circ x = x$. Hence the stabilizer is a subgroup. □

6.3 Orbit-Stabilizer Theorem

Consider a point $x \in X$. Its orbit is the set of points reachable through the action of all group elements $a \in G$. If all action points $a \circ x$ are distinct, only $e \circ x = x$ fixes x, hence its stabilizer consists of the group identity e only, the smallest possible subgroup. An example of such a group action is shown in Figure 3.

What happens if not all action points are distinct? This is interesting:

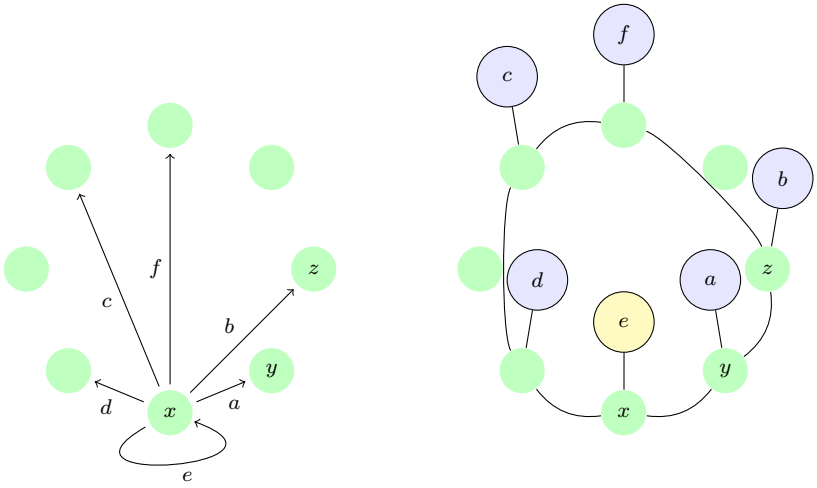

Fig. 3. Let $G = \{a, b, c, d, e, f\}$ with e being the identity, and $X = \{x, y, z, \dots\}$. If a group action, shown on the left, maps a point $x \in X$ to distinct reachable points, each reachable point corresponds to only one element in G. These reachable points can be joined together by arcs, shown on the right, giving the $orbit\ x$. The "balloon" over each $y \in orbit\ x$ contains the group element which acts on x to reach y. Note the balloon over x is $stabilizer\ x$, in this case just $\{e\}$, corresponds to the self-loop over x on the left.

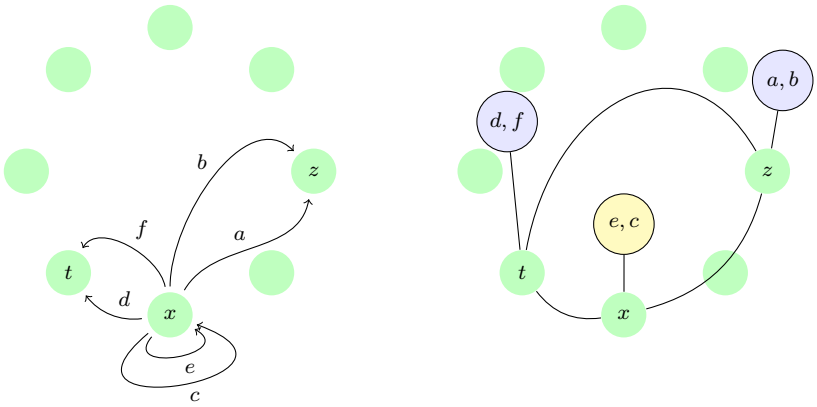

Fig. 4. If $stabilizer\ x$, shown on the left as self-loops over x, has two elements $\{e, c\}$, then every $z \in orbit\ x$ is reachable by two group elements: $z = a \circ x = a \circ (c \circ x) = (a \times c) \circ x = b \circ x$ where $b = a \times c$, and $b \neq a$ since $c \neq e$. On the right are shown the two group elements for every point in $orbit\ x$ inside its "balloon". The balloon over x is $stabilizer\ x$; the other balloons are cosets of $stabilizer\ x$.

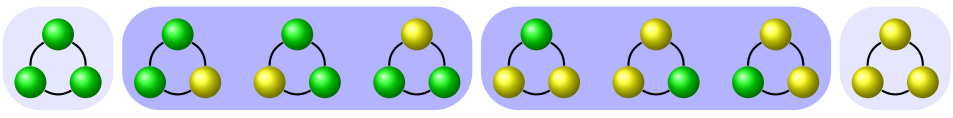

Fig. 5. Orbits of necklaces with 3 beads in 2 colours under cycle action by $\mathbb{Z}_3^+$ are shown as background round rectangles. Light blue orbits of monocoloured ones always have size $= 1$. Dark blue orbits of multicoloured ones always have size > 1. By Orbit-Stabilizer theorem, orbit size divides $|\mathbb{Z}_3^+| = 3$, hence multicoloured orbit size $= 3$.

THEOREM 15. *If action points coincide: $a \circ x = b \circ x$, the quotient $a^{-1} \times b$ lies in the stabilizer of x.*

```
⊢ Group g ∧ action (∘) g X ∧ x ∈ X ⇒
    ∀ a b::(G). a ∘ x = b ∘ x ⟺ a⁻¹ × b ∈ stabilizer x
```

PROOF. From left to right, apply the action a^{-1} to both sides of $a \circ x = b \circ x$. Thus $x = (a^{-1} \times b) \circ x$. From right to left, we have $(a^{-1} \times b) \circ x = x$. Apply the action a to both sides of this equation, deriving $b \circ x = a \circ x$ as required. □

Furthermore, $a^{-1} \times b$ is some $c \in G$ by closure property of group. By uniqueness of group inverses, $c \neq e$ if $a \neq b$. Hence for distinct $a, b \in G$ with $a \circ x = z = b \circ x$, there is an element $c \neq e$ and $c \in stabilizer\ x$. Note that $a^{-1} \times b = c$ implies $b = a \times c$. So if, for example (as shown in Figure 4), $stabilizer\ x$ is indeed just $\{e, c\}$, then the set of group elements enabling x to reach z, *i.e.* $\{a, b\} = \{a \times e, a \times c\} = a\ \{e, c\}$, is a coset of $stabilizer\ x$.

Similar reasoning shows that, for any point $y \in orbit\ x$, the set of group elements $\{a \in G \mid a \circ x = y\}$ that enables x to reach y will be a coset of $stabilizer\ x$. In HOL, this is expressed as:

THEOREM 16. *The set of group elements enabling x to reach point $y \in orbit\ x$ is a coset of $stabilizer\ x$.*

```
⊢ Group g ∧ action (∘) g X ∧ x ∈ X ∧ y ∈ orbit x ⇒
    { a | a ∈ G ∧ a ∘ x = y } =
      coset g (stabilizer x) (actionElement (∘) g x y)
```

where (`actionElement (∘) g x y`) *is a group element that acts on x, generating y.* [6]

Such an element exists when x and y are in the same orbit, as assumed. Since the choice made by `actionElement` may not be drawing on a singleton set, this association of a point $y \in orbit\ x$ with a coset is only meaningful if the coset is independent of this choice. This follows from a standard result about subgroup cosets:

THEOREM 17. *Two cosets of a subgroup are equal when it has the quotient of their generating elements.*

```
⊢ Group g ∧ h ≤ g ⇒
    ∀ b a::(G). coset g H b = coset g H a ⟺ a⁻¹ × b ∈ H
```

PROOF. For the if-part ($\Rightarrow$), since $b \in bH$, $bH = aH$ implies there is a $c \in H$ such that $b = a \times c$. Solving for c in a group: $c = a^{-1} \times b \in H$. For the only-if part ($\Leftarrow$), since $(a^{-1} \times b) \in H$, so for any $c \in H$, $(a^{-1} \times b) \times c$ equals to some $d \in H$, by closure property of a subgroup. Now $(a^{-1} \times b) \times c = d$ implies $b \times c = a \times d$, for any $c \in H$. This shows $bH \subseteq aH$. Repeating the same argument with $b^{-1} \times a = (a^{-1} \times b)^{-1} \in H$, as a subgroup includes all inverses, gives $aH \subseteq bH$. Thus $bH = aH$. □

This matching condition is used to prove the association of stabilizer cosets to orbit points (Theorem 16). Together with the matching condition of reachable points (Theorem 15), both are crucial in establishing:

THEOREM 18. *The points of x's orbit are in bijection with the cosets of x's stabilizer.*

[6] This seemingly clumsy formulation is intended to bring out the issue about choice, which is discussed subsequently and helps to compare Theorem 15 with Theorem 17. A direct formulation of the same result is:

```
⊢ Group g ∧ action (∘) g X ∧ x ∈ X ∧ y ∈ orbit x ⇒ ∀ b. b ∈ G ∧ b ∘ x = y ⇒ { a | a ∈ G
∧ a ∘ x = y } = coset g (stabilizer x) b
```

```
⊢ Group g ∧ action (∘) g X ∧ x ∈ X ⇒
    BIJ (λz. coset g (stabilizer x) (actionElement (∘) g x z))
      (orbit x) {coset g (stabilizer x) a | a | a ∈ G}
```

The last set comprehension is a special form marking a as the only variable that varies in the leftmost expression, `coset g (stabilizer x) a`. This bijection provides the key for:

THEOREM 19. *Orbit-Stabilizer Theorem.*

```
⊢ FiniteGroup g ∧ action (∘) g X ∧ x ∈ X ∧ FINITE X ⇒
    |G| = |orbit x| × |stabilizer x|
```

PROOF. There are $|orbit\ x|$ points in x's orbit. Each point is associated with a coset of $stabilizer\ x$. Since $stabilizer\ x$ is a subgroup (Theorem 14), each coset is the same size as $stabilizer\ x$ (Theorem 12). The cosets form a partition of the carrier set G (Theorem 10), which is counted by the bijection of Theorem 18:

$$|G| = \sum_{a \in G} |a(stabilizer\ x)| = |orbit\ x|\,|stabilizer\ x|$$

□

6.4 Applying Actions to Necklaces

The Orbit-Stabilizer theorem is the key to classifying necklace orbits, especially when the necklace length is prime. First we identify the group action:

THEOREM 20. `cycle` *is an action from* $\mathbb{Z}_n^+$ *to the set of necklaces:*

```
⊢ 0 < n ∧ 0 < a ⇒ action cycle ℤ_n^+ (necklace n a)
```

PROOF. For necklace $\ell \in$ `necklace n a`, $|\ell| =$ `n`. Each element $k \in \mathbb{Z}_n^+$, *i.e.* $0 \leq k < n$, maps a necklace ℓ to the cycle result: `cycle k ℓ`, *i.e.* cycling of the necklace by k beads. Recall these earlier results about `cycle` (Section 3.2):

```
⊢ ℓ ∈ necklace n a ⇒ ∀k. cycle k ℓ ∈ necklace n a
⊢ cycle 0 ℓ = ℓ
⊢ ℓ ≠ [] ⇒ cycle x (cycle y ℓ) = cycle ((x + y) mod |ℓ|) ℓ
```

The first shows `cycle` is closed for necklaces. The second shows `cycle` has an identity. The third shows `cycle` composes under modulo n addition. Hence `cycle` is an action from the group $\mathbb{Z}_n^+$. □

Since length and colours are invariants for `cycle` (Section 3.2), a multicoloured necklace cannot be cycled to a monocoloured necklace. This shows `cycle` is also closed for those sets respectively:

```
⊢ 0 < n ∧ 0 < a ⇒ action cycle ℤ_n^+ (monocoloured n a)
⊢ 0 < n ∧ 0 < a ⇒ action cycle ℤ_n^+ (multicoloured n a)
```

The classification of orbits for necklaces is simple:

THEOREM 21. *Only monocoloured necklaces have orbit size equal to* 1.

```
⊢ 0 < n ∧ 0 < a ∧ ℓ ∈ monocoloured n a ⇒
    |orbit cycle ℤ_n^+ (monocoloured n a) ℓ| = 1
⊢ 0 < n ∧ 0 < a ∧ ℓ ∈ multicoloured n a ⇒
    |orbit cycle ℤ_n^+ (multicoloured n a) ℓ| ≠ 1
```

PROOF. Only a monocoloured necklace ℓ has `cycle 1 ℓ = ℓ` (Theorem 2), *i.e.* for all multiples k, `cycle k ℓ = ℓ` by `CYCLE_ADD`. Hence only such orbit collapses to a singleton, with cardinality 1. □

THEOREM 22. *For multicoloured necklaces of length p, a prime, orbit size of each necklace equals p.*

$$\vdash \texttt{prime}\ p \wedge 0 < a \wedge \ell \in \texttt{multicoloured}\ p\ a \Rightarrow |\texttt{orbit cycle}\ \mathbb{Z}_p^+\ (\texttt{multicoloured}\ p\ a)\ \ell| = p$$

PROOF. When the necklace length is prime p, the action group is $\mathbb{Z}_p^+$. By the Orbit-Stabilizer theorem (Theorem 18): $|orbit\ \ell| \times |stabilizer\ \ell| = |\mathbb{Z}_p^+| = p$ for any necklace ℓ. A prime p has only trivial factorization: $p = 1 \times p = p \times 1$. By Theorem 21, the orbit of a multicoloured necklace is not a singleton, so its size must be p. □

Recall that `reach` is an equivalence relation (Theorem 13). Orbits are the equivalence classes of `reach`, so they form a partition of the target set. From Theorem 22, the target is (`multicoloured` p a) with size $(a^p - a)$ (Section 3.1). Theorem 22 also specifies an equal-size partition, giving the visual image of division (Figure 5 and Figure 6), which can be described as *visual divisibility*, see next Section 6.5. Hence, p `divides` $a^p - a$, which in modulo p is Fermat's Little Theorem:

$$\vdash \texttt{prime}\ p \Rightarrow a^p \equiv a\ (\mathrm{mod}\ p)$$

The actual HOL4 proof of this assertion using the aforementioned reasoning is given in Appendix.

6.5 Insights from Group Actions

Group actions give rise to orbits, which are equivalence classes of `reach` (Section 6.2). In general, equivalence classes form a partition. When the equivalence classes are all of the same size, this is an equal-size partition — or *visual divisibility*. Applying to action orbits, this means:

THEOREM 23. *If all action orbits are all of the same size, the orbit size divides the size of target set.*

$$\vdash \texttt{Group}\ g \wedge \texttt{action}\ (\circ)\ g\ X \wedge \texttt{FINITE}\ X \wedge (\forall x.\ x \in X \Rightarrow |orbit\ x| = n) \Rightarrow n\ \texttt{divides}\ |X|$$

PROOF. As a consequence of Theorem 13, action orbits form a partition of the target set X. Since a partition is both exclusive and exhaustive:

$$|X| = \sum_{x \in X} |orbit\ x|$$

When all orbits are of the same size n, *i.e.* $|orbit\ x| = n$, we have:

$$|X| = k \times n$$

where k is the number of distinct orbits. The divisibility result follows. □

Petersen's proof (Section 3) in essence identifies necklace cycle permutations with action orbits. The heart of his short proof is the following claim:

THEOREM 24. Necklace Theorem
The necklace orbit size is always a factor of the number of necklace beads.

$$\vdash 0 < n \wedge 0 < a \Rightarrow \forall \ell.\ \ell \in \texttt{necklace}\ n\ a \Rightarrow |\texttt{orbit cycle}\ \mathbb{Z}_n^+\ (\texttt{necklace}\ n\ a)\ \ell|\ \texttt{divides}\ n$$

PROOF. Since `cycle` is an action from the group $g = \mathbb{Z}_n^+$ to the target set $X =$ `necklace` n a (Theorem 20), this is a direct consequence of the Orbit-Stabilizer formula: $|\mathtt{G}| = |orbit\ x| \times |stabilizer\ x|$ from Theorem 19. □

Fermat's Little Theorem is a simple corollary of this theorem, because primes have only trivial factors. Indeed, in this case it gives equal-size partitions for multicoloured necklace orbits, providing a direct manifestation of visual divisibility (Theorem 23).

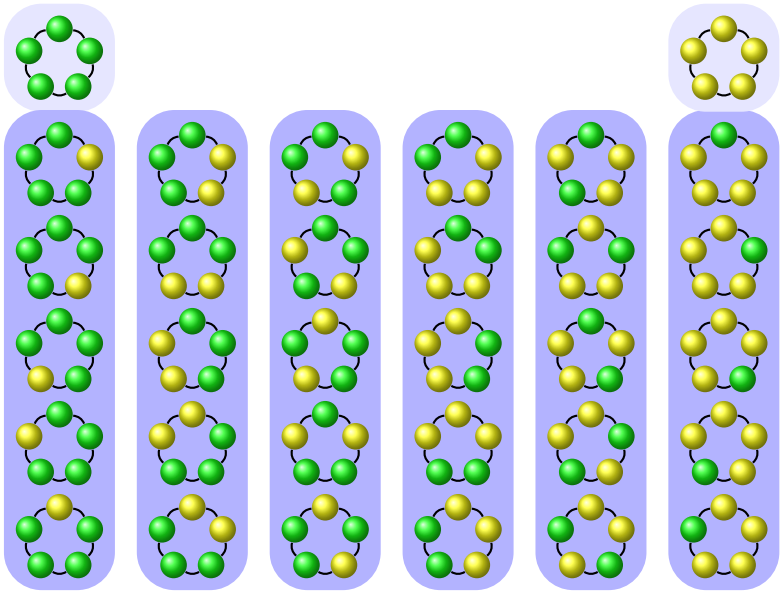

Fig. 6. Orbits of necklaces with 5 beads in 2 colours under cycle action by $\mathbb{Z}_5^+$. The monocoloured necklaces (top left and top right) each has orbit of size 1. The multicoloured necklaces have orbits in a regular pattern, in sets of size 5. By Orbit-Stabilizer theorem, the possible orbit sizes of necklaces are 1 or 5, since $n = 5$ is prime.

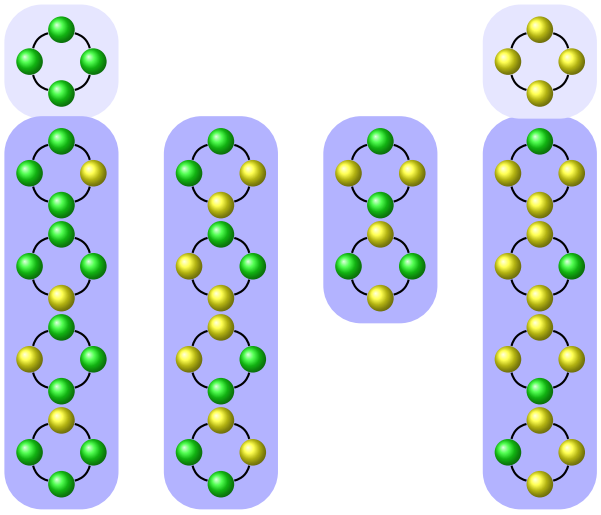

Fig. 7. Orbits of necklaces with 4 beads in 2 colours under cycle action by $\mathbb{Z}_2^+$ The monocoloured necklaces (top left and top right) each has orbit of size 1. The multicoloured necklaces have orbits in an irregular pattern, with sets of size 4 or 2. By Orbit-Stabilizer theorem, the possible orbit sizes of necklaces are divisors of 4, *i.e.* 1, 2 or 4.

Petersen regarded *Necklace Theorem* as "obvious", and his three-line statement is just a proof of this corollary (compare the proofs of Theorem 23 and Theorem 6). However, if one is pressed for details, as is demanded for any formalization, a proof of this *Necklace Theorem* is required – as we have done *via* Orbit-Stabilizer Theorem.

Visual patterns of necklace orbits from *Necklace Theorem* are given in Figure 5, Figure 6 and Figure 7, which are graphical depictions of visual divisibility (Theorem 23). These pictures of *Necklace Theorem* explain why the converse of Fermat's Little Theorem is plausible, but not necessarily so. When the number of beads is not a prime, the partition by multicoloured necklace orbits is not of equal size. While divisibility is not likely for an irregular-sized partition, it may happen in exceptional cases. Indeed, those rare composite numbers n that satisfy $a^n \equiv a \pmod{n}$ for any a are known as *Carmichael numbers*.

7. ACTION FIXED-POINTS

We have seen how an action from a group G to a target set X gives rise to orbits (points moved by group elements) and stabilizers (group elements fixing a point). Indeed, those special points in the target set that are fixed by all group elements form a special set with a specific name:

DEFINITION 6. *The **fixed points** of an action is the set of target points fixed by the group:*

$\vdash$ `fixedpoints (∘) g X = {x | x ∈ X ∧ ∀a::(G). a ∘ x = x}`

Thus each fixed point has the whole group as its stabilizer, or equivalently, its orbit must be a singleton by Orbit-Stabilizer Theorem (Theorem 19):

```
⊢ Group g ∧ action (∘) g X ⇒
    ∀a. a ∈ X ⇒ (a ∈ fixedpoints (∘) g X ⟺ SING (orbit a))
```

Consider an action from the group $\mathbb{Z}_p^+$ to a finite target set X, where p is prime. We have:

THEOREM 25. *For action of $\mathbb{Z}_p^+$ with prime p, the target size is congruent to fixed points size in modulo p.*

$\vdash$ `prime p ∧ action (∘) ℤ_p^+ X ∧ FINITE X ⇒`
`|X| ≡ |fixedpoints (∘) ℤ_p^+ X| (mod p)`

PROOF. Since p is prime, the non-singleton orbits are all of size p by Orbit-Stabilizer Theorem 19. These orbits form an equal-size partition within the set NF = `X \ fixedpoints (∘) ℤ_p^+ X` of all non-fixed points, *i.e.* $|NF|$ is a multiple of p. Since the sets fixed points and non-fixed points are disjoint, $|NF|$ = $|X|-|$`fixedpoints (∘) ℤ_p^+ X`$|$, and the congruence equality follows. □

Applying this action to necklaces (`necklace n a`), and identifying the `monocoloured` necklaces as fixed points of this action (recall that only monocoloured necklaces have singleton orbits, by Theorem 21):

THEOREM 26. *The fixedpoints of* `cycle` *are monocoloured necklaces.*

```
⊢ 0 < n ∧ 0 < a ⇒
    fixedpoints cycle ℤ_n^+ (necklace n a) = monocoloured n a
```

We arrive at another demonstration of

THEOREM 27. *Fermat's Little Theorem by Fixed Points.*

$\vdash$ `prime p ⇒` $a^p \equiv a \pmod{p}$

PROOF. By Theorem 20, we know `cycle` is an action from $\mathbb{Z}_p^+$ to `necklace p a`. For this action, the fixed points are just the `monocoloured` necklaces, as shown by Theorem 26. From Section 3.1, the target set $|X| = |$`necklace p a`$| = a^p$, and $|$`fixedpoints cycle ℤ_p^+ (necklace p a)`$| =$ $|$`monocoloured p a`$| = a$. Thus the result follows from the congruence in Theorem 25. □

8. GROUP THEORY APPLIED TO THE NUMBER-THEORETIC PROOF

Having applied group theory to the necklace proof, it is interesting to try the "same trick" with the number-theoretic proof. The subsequent results are not novel, but allow a fuller comparison of approaches when we conclude.

It is straightforward to recast the number-theoretic proof of Fermat's Little Theorem (Section 4) in the context of finite Abelian groups, the structure that naturally mimics prime modulo multiplication. The factor rearrangement and cancellation are direct consequences of commutativity and cancellation laws in Abelian groups. However, this is not very illuminating, and unnecessarily restrictive, as the group-theoretic version of Fermat's Little Theorem holds for all finite groups (not just the Abelian ones):

$$\vdash \texttt{FiniteGroup}\ g \wedge a \in \texttt{G} \Rightarrow a^{|\texttt{G}|} = e$$

Assuming this result (which will be proved later, see Theorem 29), to derive Fermat's Little Theorem (*i.e.* $a^{p-1} \equiv 1 \pmod{p}$) it is sufficient to deomonstrate that prime modulo multiplication, *i.e.* $\mathbb{Z}_p^*$ for prime p, does indeed form a group — with $|\mathbb{Z}_p^*| = |\{1 \ldots p-1\}| = p-1$, and the multiplicative identity is 1. Critically, we need to show that any $x \in \{1 \ldots p-1\}$ has a multiplicative inverse in $\mathbb{Z}_p^*$. This can be done by appeal to Bézout's identity, a property of `gcd` already used in Theorem 1:

$$\vdash x \neq 0 \Rightarrow \exists k\ q.\ k \times x = q \times p + \texttt{gcd}\ p\ x$$

With p a prime and $0 < x < p$, $\texttt{gcd}\ p\ x = 1$. Taking modulo p on both sides of the equation, the right-hand side becomes 1, and the k on the left gives $k \pmod{p}$ as the multiplicative inverse of x in $\mathbb{Z}_p^*$.

8.1 Euler's Generalisation

When the modulo n is not a prime, the non-zeroes of $\mathbb{Z}_n$ do not form a multiplicative group; *e.g.* to find the multiplicative inverse for 2 in $\mathbb{Z}_6$ would require solving $2x = 6y + 1$, which is impossible by parity. However, Euler observed that the Bézout's identity of the preceding section actually guarantees a multiplicative inverse for $0 < x < n$ whenever $\texttt{gcd}\ n\ x = 1$, *i.e.* x is co-prime to n:

$$\begin{array}{l}\vdash 1 < n \wedge 0 < x \wedge x < n \wedge \texttt{coprime}\ n\ x \Rightarrow \\ \quad \exists y.\ 0 < y \wedge y < n \wedge \texttt{coprime}\ n\ y \wedge y \times x \bmod n = 1\end{array}$$

This is then the basis for a group, whose carrier is $\mathbb{Z}_n^*$ — the set of elements of $\mathbb{Z}_n$ with multiplicative inverses. The cardinality of this set is known as its *totient*, denoted by $\varphi(n)$:

$$\begin{array}{l}\vdash \mathbb{Z}_n^* = \{x \mid 0 < x \wedge x < n \wedge \texttt{coprime}\ n\ x\} \\ \vdash \varphi(n) = |\mathbb{Z}_n^*|\end{array}$$

With all these at hand, this is evident:

THEOREM 28. *Euler's generalisation of Fermat's Little Theorem*

$$\vdash 1 < n \wedge 0 < a \wedge a < n \wedge \texttt{coprime}\ n\ a \Rightarrow a^{\varphi(n)} \equiv 1 \pmod{n}$$

PROOF. By Theorem 29 below. □

We shall now prove the group-theoretic version of Fermat's Little Theorem for any finite group, *via* the generated subgroup of its elements.

8.2 Generated subgroups

Let a be a group element. Consider the sequence of powers: $a, a^2, a^3, \ldots$. If the group is finite, there must eventually be a repetition in this sequence. Assume $m < n$ and $a^m = a^n$, then we can use left-cancellation in group to remove the common factor a^m, giving us

$$\vdash \texttt{Group}\ g \wedge a \in \texttt{G} \wedge m < n \wedge a^m = a^n \Rightarrow a^{n-m} = e$$

DEFINITION 7. *Call the least non-zero exponent that maps an element back to the identity, its* ***order****:*

$\vdash$ `order` $g\ a$ = `LEAST` k. `0` $< k \wedge a^k = e$

The preceding argument shows that `order` exists for finite group elements, and it satisfies:

$\vdash$ `FiniteGroup` $g \wedge a \in$ `G` $\Rightarrow$ `0` $<$ `order` $g\ a \wedge a^{\texttt{order}\ g\ a} = e$
$\vdash$ `FiniteGroup` $g \wedge a \in$ `G` $\Rightarrow a^{-1} = a^{\texttt{order}\ g\ a-1}$

By properties of group exponentiation (Section 5), the powers of an element $a \in G$ form a subgroup: `Generated` $g\ a$, also written as $\langle a \rangle$. This subgroup is related to `order` by:

$\vdash$ `FiniteGroup` $g \wedge a \in$ `G` $\Rightarrow$ $|$`(Generated` $g\ a$`).carrier`$|$ = `order` $g\ a$

This result can be deduced by the `LEAST` property of `order`, and provides the key for:

THEOREM 29. *Fermat's Little Theorem for finite groups.*

$\vdash$ `FiniteGroup` $g \wedge a \in$ `G` $\Rightarrow a^{|\texttt{G}|} = e$

PROOF. Consider $\langle a \rangle$, the generated subgroup of $a \in$ `G`, with cardinality $|\langle a \rangle| =$ `order` $g\ a$ for a finite group. By definition of `order`, $a^{\texttt{order}\ g\ a} = e$. Since $\langle a \rangle$ is a subgroup of g, by Lagrange's Theorem there exists k such that $|\texttt{G}| = |\langle a \rangle| \times k =$ `order` $g\ a \times k$. Hence,

$$a^{|\texttt{G}|} = a^{\texttt{order}\ g\ a \times k} = (a^{\texttt{order}\ g\ a})^k = e^k = e$$

as required. □

9. INDUCTION USING BINOMIALS

Most likely the earliest proof of Fermat's Little Theorem is given by induction via the binomial theorem (Section 2), since the coefficients of binomial expansion form the well-known "Pascal's triangle". In HOL4 we can establish[7]:

DEFINITION 8. *The* ***binomial coefficients*** *are defined recursively:*

$$\binom{n}{0} = 1$$
$$\binom{0}{k+1} = 0$$
$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$$

For a given integer n, its unit binomials are: $\binom{n}{0} = 1$ and $\binom{n}{n} = 1$. All binomial coefficients are integers, and there is a formula to compute them directly *via* factorials:

$$\vdash k < n \Rightarrow \binom{n}{k} = \frac{n!}{k! \times (n-k)!}$$

The standard binomial expansion with exponent n can be expressed as a sum over terms with binomial coefficients:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$$

In HOL4 this is expressed as:

[7]The HOL4 source code provides an example of proving Fermat's Little Theorem using the Binomial Theorem. This proof is by Laurent Théry, and is apparently itself a translation of a Coq proof by J. C. Almeida.

$$\vdash \texttt{(}x \texttt{ + } y\texttt{)}^{n} \texttt{ = SUM (GENLIST } (\lambda k.\ \binom{n}{k} \times x^{n-k} \times y^{k}) \texttt{ (}n \texttt{ + 1))}$$

where SUM performs a summation over GENLIST f (n + 1), a list generated by applying function f to the index list $[0; \ldots ; n]$ with (n + 1) terms.

The binomial coefficients for a prime have a special property:

THEOREM 30. *A prime divides all its non-unit binomials.*

$$\vdash \texttt{prime } p \Rightarrow 1 < p \wedge \forall k.\ 0 < k \wedge k < p \Rightarrow p \texttt{ divides } \binom{p}{k}$$

PROOF. Rearrange the binomial formula to: $k! \times \texttt{(}p-k\texttt{)}! \times \binom{p}{k} = p!$. Clearly p divides $p!$, the right-hand side, so p must also divide the left-hand side. Now prime p cannot divide $k!$ or $\texttt{(}p-k\texttt{)}!$ for $0 < k < p$, since these factorials are composed of numbers all less than p. Therefore prime p must divide $\binom{p}{k}$ for all such k's, which are all its non-unit binomials. □

Due to this crucial property of primes, the binomial expansion involving prime exponent and modulo takes a simple form:

THEOREM 31. *When both exponent and modulo are the same prime, the binomial expansion under modulo has the exponent distributing over the binomial.*

$$\vdash \texttt{prime } p \Rightarrow \forall x\ y.\ \texttt{(}x \texttt{ + } y\texttt{)}^{p} \equiv x^{p} \texttt{ + } y^{p} \pmod{p}$$

PROOF. Note that, by Theorem 30, for prime p and $0 < k < p$ we have $\binom{p}{k} \equiv 0 \pmod{p}$ since perfect division leaves no remainder. Therefore,

$$\begin{aligned}
& (x+y)^p \pmod{p} \\
\equiv\ & \sum_{k=0}^{n} \binom{p}{k} x^{p-k} y^{k} \pmod{p} && \text{by binomial theorem} \\
\equiv\ & \binom{p}{0} x^{p} + \sum_{k=1}^{n-1} \binom{p}{k} x^{p-k} y^{k} + \binom{p}{p} y^{p} \pmod{p} && \text{by breaking out first and last terms} \\
\equiv\ & 1 \times x^{p} + \sum_{k=1}^{n-1} 0 \times x^{p-k} y^{k} + 1 \times y^{p} \pmod{p} && \text{by Theorem 30 and unit binomials} \\
\equiv\ & x^{p} + y^{p} \pmod{p}
\end{aligned}$$

□

This leads to another proof of Fermat's Little Theorem (probably Leibniz's proof, see Section 2):

THEOREM 32. *Fermat's Little Theorem via Binomial Theorem.*

$$\vdash \texttt{prime } p \Rightarrow a^{p} \equiv a \pmod{p}$$

PROOF. By induction on the base a, with fixed exponent the prime p. For $a = 0$, the congruence identity is trivially true. Assume that the congruence identity holds for some base a: *i.e.* $a^{p} \equiv a \pmod{p}$, then for the next base $(a+1)$:

$$\begin{aligned}
& \texttt{(}a \texttt{ + 1)}^{p} \pmod{p} \\
\equiv\ & a^{p} \texttt{ + 1}^{p} \pmod{p} && \text{by Theorem 31} \\
\equiv\ & a^{p} \bmod p \texttt{ + 1} \bmod p \pmod{p} && \text{by modulo addition} \\
\equiv\ & a \bmod p \texttt{ + 1} \bmod p \pmod{p} && \text{by induction hypothesis} \\
\equiv\ & a \texttt{ + 1} \pmod{p} && \text{by modulo addition again}
\end{aligned}$$

□

10. CONCLUSION

Fermat's Little Theorem is a very basic and well-known result in number theory. Having attempted its proof in a slew of different styles, we now attempt to draw some lessons.

Analysis. For each proof discussed, we have linearised our various script files into one script containing just the lemmas required for that particular effort. Table I includes total line counts[8] for each file.

Type of Proof	Approach (Section reference)		Filename	Total
Combinatorial	Direct *via* cycles	(3)	`AllFLTnecklaceScript.sml`	766
	Group *via* action	(6)	`AllFLTactionScript.sml`	1348
	Group *via* fixed points	(7)	`AllFLTfixedpointsScript.sml`	1529
Number-theoretic	Direct *via* modulo arithmetic	(4)	`AllFLTnumberScript.sml`	400
	Group *via* generated subgroups	(8)	`AllFLTgroupScript.sml`	787
	Euler *via* generated subgroups	(8)	`AllFLTeulerScript.sml`	819
	Induction *via* binomial theorem	(9)	`AllFLTbinomialScript.sml`	597

Table I. Line counts for theories developing each approach. Filename is that of the linearised script in `bitbucket.org` repository.

Type of Proof	Approach (Section reference)		Basis	General Theory	Actual Proof	Total
Combinatorial	Direct *via* cycles	(3)	126	359	281	766
	Group *via* action	(6)	167	1030	151	1348
	Group *via* fixed points	(7)	181	1185	163	1529
Number-theoretic	Direct *via* modulo arithmetic	(4)	129	0	271	400
	Group *via* generated subgroups	(8)	32	604	151	787
	Euler *via* generated subgroups	(8)	32	604	183	819
	Induction *via* binomial theorem	(9)	50	356	191	597

Table II. Each linearized proof script consists of 3 sections: basis, general theory and actual proof. Individual line counts of each section, making up the total, are shown.

In order to analyze the total line counts further, we split each linearized script into three sections:

— *Basis*: proofs of minor properties about lists, numbers and sets not already present in the standard HOL4 distribution, but required for this mechanisation. These results arguably represent holes in HOL4's general library of facts about these core types. More generously, results proved here are just re-expressions of theorems already in the system, but tailored to suit our particular purpose.

[8] A suggestion to consider the de Bruijn factor, whereby one compares the zipped size of the proof scripts, is not adopted here since we are interested in comparing the *relative* size of proof scripts, written by a single author on a single system.

— *General theory*: results about the general theory behind the proof. In the case of the necklace proofs, results included here are properties of the `cycle` construct, and the categorisation into mono- and multi-coloured necklaces. In the case of proofs depending on group theory, this section includes all of the necessary group theory and further developments (e.g. group action theory and action fixed points) for the given result. In the case of the induction proof using binomials, this section includes definitions and properties of binomial coefficients. There is no additional general theory required for the direct number-theoretic proof.

— *Actual proof*: the development of the argument leading to the final result. For example, in the direct necklace proof, this includes the development of the similarity equivalence relation, and the notion of `order`. In the group action proof, this is just the application of Orbit-Stabiliser theorem to the case of multicoloured necklaces. In the proof by action fixed points, this includes the identification of `cycle` action fixed points with the monocolured necklaces. In the number-theoretic proof based on group theory, the actual proof includes the result that $\mathbb{Z}_p^*$ really is a group, but the proof of the group-theoretic analogue of Fermat's Little Theorem (Theorem 29) is in the "general theory". Clearly then, there are judgement calls to be made here about whether or not a particular theorem should be in the "general theory" or in the "actual proof".

Table II includes the detailed counts for each proof approach.

The verdict is clear: the basic number-theoretic approach is much better in terms of overall lines-of-code. Also somewhat dispiriting is the fact that the combinatorial necklace proofs are all worse than all the number-theoretic proofs by the same metric. Though the necklace proofs might perform better if HOL4's library of facts about lists were better tailored to their needs, it's also undeniable that the general theory of list-rotations is unlikely to be as useful as the general group theory used in the number-theoretic proofs. The only real solace for the necklace proof (recall Section 3: just three lines of English prose!), is that when the independently interesting group theory of actions and orbits is to hand, the actual proof is among the best — indeed by using group theory, the action-based group proof of Section 6.4 and number-theoretic group proof of Section 8 both score the shortest actual proof.

However, these results can only be suggestive: our HOL scripts may not be the optimal expression of the various proof strategies, and our own skills may not be uniform across the numerical and algebraic domains. HOL4 supports different proof styles, but the authors favour the use of assertion style that provides an easy-to-follow train of reasoning, *i.e.* the HOL proof in Appendix.

Of course, HOL4's features make some proofs easier to automate, and some goals easier to express. Certainly, we believe that our naïve approach to group theory makes the numbers for the group-theoretic proofs worse than they might be, and the Orbit-Stabiliser theorem is arguably a steeper requirement than Fermat's Little Theorem for finite groups (Theorem 29). Dependent types, perhaps best exemplified by their use in Coq (though also approximated and used for group theory in HOL4 by Hurd [Hur01]), would be an obvious way to approach this issue. For example, it is painful to first quantify over the record type of four fields from Section 5, and then have to additionally assume that the structure in question really satisfies the group axioms (using the `Group` predicate). Isabelle's axiomatic type-classes and locales have also been used to provide appealing mechanisations of abstract algebra. It would be interesting to see what these other systems made of the directly combinatorial necklace proof, and of the group-theoretic version of the same.

Future Work. Fermat's Little Theorem and Binomial Theorem are crucial concepts in Agrawal *et al.*'s famous result [AKS04] that primality testing can be done in polynomial time. We are working towards a mechanisation of this result.

Final Verdict. Our results show that we have yet to find the sweet spot when it comes to performing combinatorial proofs in HOL. Our consolation is to have found that attacking the result *via* explicit appeals to group theory gives us two distinct mechanised proofs that are arguably more elegant than their

direct analogues. Our mechanisation of the necklace proofs may be the first; we hope it is not the last, and that still more beautiful pearls may be found in this vein.

Appendix

Here is the HOL4 proof of Fermat's Little Theorem based on necklaces and group action, discussed in Section 6.4. This proof is given at the end of `AllFLTactionScript.sml`, with comments marked by `(* ... *)`. It refers to various established theorems by name:

— `ZN_group` asserts that the group definition applies to $\mathbb{Z}_n^+$ introduced in Section 6:

$\vdash$ `0 < n ⇒ Group` $\mathbb{Z}_n^+$

— `CARD_multicoloured` is given in Section 3.1:

$\vdash$ `0 < n ⇒ |multicoloured n a| =` $a^n - a$

— `cycle_action_on_multicoloured` is given in Section 6.4:

$\vdash$ `0 < n ∧ 0 < a ⇒ action cycle` $\mathbb{Z}_n^+$ `(multicoloured n a)`

— `CARD_multicoloured_PRIME_ORBIT` is proved as Theorem 22:

$\vdash$ `prime p ∧ 0 < a ∧ ℓ ∈ multicoloured p a ⇒`
`|orbit cycle` $\mathbb{Z}_p^+$ `(multicoloured p a) ℓ| = p`

— `EQUAL_SIZE_ORBITS_PROPERTY` is proved as Theorem 23:

$\vdash$ `Group g ∧ action (∘) g X ∧ FINITE X ∧ (∀x. x ∈ X ⇒ |orbit x| = n) ⇒`
`n divides |X|`

— `PRIME_POS`, `SUB_0`, `EXP_EQ_0`, and `ALL_DIVIDES_0` are provided in HOL4 libraries.

PROOF. The HOL4 proof proceeds in a fairly readable assertion style, as shown.

```
val FERMAT_LITTLE_THM = store_thm(
  "FERMAT_LITTLE_THM",
  ``!p a. prime p ==> divides p (a**p - a)``,
  REPEAT STRIP_TAC THEN
  `0 < p` by RW_TAC std_ss [PRIME_POS] THEN
  Cases_on `a = 0` THENL [
    (* case: (a = 0) /\ prime p ==> divides p (a ** p - a) *)
    METIS_TAC [SUB_0, EXP_EQ_0, ALL_DIVIDES_0],
    (* case: a <> 0 /\ prime p ==> divides p (a ** p - a) *)
    `0 < a` by DECIDE_TAC THEN
    `CARD (multicoloured p a) = a**p - a`
       by RW_TAC std_ss [CARD_multicoloured] THEN
    `Group (Z p)` by RW_TAC std_ss [ZN_group] THEN
    `action cycle (Z p) (multicoloured p a)`
       by RW_TAC std_ss [cycle_action_on_multicoloured] THEN
    `FINITE (multicoloured p a)`
       by RW_TAC std_ss [FINITE_multicoloured] THEN
    `!l. l IN (multicoloured p a) ==>
      (CARD (orbit cycle (Z p) (multicoloured p a) l) = p)`
       by RW_TAC std_ss [CARD_multicoloured_PRIME_ORBIT] THEN
    METIS_TAC [EQUAL_SIZE_ORBITS_PROPERTY]
  ]);
```

□

References

[AA08] Andrea Asperti and Cristian Armentano. A page in number theory. *Journal of Formal Reasoning*, 1(1):1–23, 2008.

[ABR05] Peter G. Anderson, Arthur T. Benjamin, and Jeremy A. Rouse. Combinatorial proofs of Fermat's, Lucas's, and Wilson's Theorems. *The American Mathematical Monthly*, 112(3):266–268, 2005.

[AKS04] Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. PRIMES is in P. *Annals of Mathematics*, 160(2):781–793, 2004.

[Bur02] Bob Burn. Fermat's Little Theorem: Proofs that Fermat might have used. *The Mathematical Gazette*, 86(507):415–422, November 2002.

[CN12] Hing-Lun Chan and Michael Norrish. A String of Pearls: Proofs of Fermat's Little Theorem. In Chris Hawblitzel and Dale Miller, editors, *Proceedings of Certified Programs and Proofs*, number 7679 in LNCS, pages 188–207. Springer, December 2012.

[Con08] Keith Conrad. Group actions. `http://www.math.uconn.edu/~kconrad/blurbs/grouptheory/gpaction.pdf`, 2008.

[Dic19] Leonard Eugene Dickson. *History of the Theory of Numbers: Volume 1: Divisibility and Primality*. Carnegie Institution of Washington, 1919.

[Gol56] Solomon W. Golomb. Combinatorial proof of Fermat's "Little" Theorem. *The American Mathematical Monthly*, 63(10):718, 1956.

[Gun89] Elsa. L. Gunter. Doing algebra in simple type theory. Technical Report MS-CIS-89-38, Department of Computer and Information Science, Moore School of Engineering,University of Pennsylvania, June 1989.

[Har11] John Harrison. *HOL Light Tutorial (for version 2.20)*. Intel JF1-13, 2011. Section 18.2: Fermat's Little Theorem.

[HE] Benjamin V. Holt and Tyler J. Evans. A group action proof of Fermat's Little Theorem. `http://arxiv.org/abs/math/0508396`.

[HGF06] Joe Hurd, Mike Gordon, and Anthony Fox. Formalized elliptic curve cryptography. In *High Confidence Software and Systems: HCSS 2006*, April 2006.

[Hur01] Joe Hurd. Predicate subtyping with predicate sets. In Richard J. Boulton and Paul B. Jackson, editors, *14th International Conference on Theorem Proving in Higher Order Logics: TPHOLs 2001*, volume 2152 of *Lecture Notes in Computer Science*, pages 265–280. Springer, September 2001.

[Mah94] Michael S. Mahoney. *The Mathematical Career of Pierre de Fermat, 1601-1665*. Princeton University Press, 1994.

[MOS06] David Marshell, Edward Odell, and Michael Starbird. *Number Theory Through Inquiry*. Mathematical Association of America Textbooks, 2006. Chapter 4: Fermat's Little Theorem and Euler's Theorem.

[Oos] Martijn Oostdijk. Library pocklington.fermat. `http://coq.inria.fr/pylons/contribs/files/Pocklington.fermat.html`.

[Rou03] Jeremy Rouse. Combinatorial proofs of congruences. Master's thesis, Harvey Mudd College, 2003.

[Rus07] David Russinoff. ACL2 Version 3.2 source: `books/quadratic-reciprocity/fermat.lisp`, 2007.

[San03] Edward Sandifer. How Euler Did It: Fermat's Little Theorem. *MAA Online*, November 2003.

[Smy86] Chris J. Smyth. A coloring proof of a generalisation of Fermat's Little Theorem. *The American Mathematical Monthly*, 93(6):469–471, 1986.

[SN08] Konrad Slind and Michael Norrish. A brief overview of HOL4. In Otmane Ait Mohamed, César Muñoz, and Sofiène Tahar, editors, *Theorem Proving in Higher Order Logics, 21st International Conference*, volume 5170 of *Lecture Notes in Computer Science*, pages 28–32. Springer, 2008.

[Wei84] André Weil. *Number Theory: An Approach Through History from Hammurapi to Legendre*. Birkhäuser Boston, 1984.

[Wik] Wikipedia: Proofs of Fermat's Little Theorem. `http://en.wikipedia.org/wiki/Proofs_of_Fermat's_little_theorem`.