

# Experiment ‘p4\_1\_a\_sum\_even\_numbers’ Results

December 18, 2025

**Experiment outcome:** SUCCESS

**Bad responses:** 0

**Responses containing assume :** 0

**Resolution attempts:** 2

**Hard fails (resolution):** 0

**Soft fails (resolution):** 0

**Verification attempts:** 2

## Problem Specification

**Problem name:** p4\_1\_a\_sum\_even\_numbers

**Natural language statement:** Write a method with loops that computes the sum of all even numbers between 2 and 100 (inclusive).

**Method signature:** p4\_1\_a\_sum\_even\_numbers() returns (sum: int)

### Ensures

- `sum == 2550`

## GenAI interactions

Below you will find all interactions between the ‘user’ (program) and the ‘assistant’ (OpenAI).

### Program → GenAI

You are given the following task to perform in Dafny:

Write a method with loops that computes the sum of all even numbers between 2 and 100 (  
↪ inclusive).

The signature should be:

```
method p4_1_a_sum_even_numbers() returns (sum: int)
```

The method should respect the following contract:

```
ensures sum == 2550
```

Produce and show only the Dafny body of this method, including the curly braces that  
↪ surround it. Do not show the signature nor contract. You must not use ‘assume’  
↪ anywhere in your code.

### GenAI → Program

**System fingerprint:**

**ID:** chatmpl-Co4ZCPohmGdckmqGjaWLaPN0gR9wm

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049990

**Finish reason:** stop

**Usage:** 260 tokens in, and 412 tokens out

```
{
    var i := 2;
    sum := 0;
    while i <= 100
        invariant 2 <= i <= 102
        invariant sum == ((i - 2) * i) / 4
    {
        sum := sum + i;
        i := i + 2;
    }
}
```

### Program → GenAI

When using dafny verify, the below error is emitted and verify fails:

```
problem_data/p4_1_a_sum_even_numbers/program.dfy(3,0): Error: a postcondition could not
    ↪ be proved on this return path
```

```

3 | {
| ^
problem_data/p4_1_a_sum_even_numbers/program.dfy(2,13): Related location: this is the
  ↪ postcondition that could not be proved
|
2 |   ensures sum == 2550
| ^^

Dafny program verifier finished with 0 verified, 1 error

Correct the error by altering only the method body. Produce and show only the Dafny body,
  ↪ including the curly braces that surround it. Do not show the signature nor
  ↪ contract. You must not use 'assume' anywhere in your code.

```

## GenAI → Program

**System fingerprint:**

**ID:** chatcmpl-Co4ZJBiNz2DYNrxZ7bILiOKguVyXB

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049997

**Finish reason:** stop

**Usage:** 581 tokens in, and 371 tokens out

```
{
  var i := 2;
  sum := 0;
  while i <= 100
    invariant i % 2 == 0
    invariant 2 <= i <= 102
    invariant sum == ((i - 2) / 2) * (((i - 2) / 2) + 1)
  {
    sum := sum + i;
    i := i + 2;
  }
}
```

## Final Program

```

method p4_1_a_sum_even_numbers() returns (sum: int)
  ensures sum == 2550
{
  var i := 2;
  sum := 0;
  while i <= 100
    invariant i % 2 == 0
    invariant 2 <= i <= 102
    invariant sum == ((i - 2) / 2) * (((i - 2) / 2) + 1)
}

```

```
    sum := sum + i;
    i := i + 2;
}
}
```

## Total Token Usage

**Input tokens:** 841

**Output tokens:** 783

**Reasoning tokens:** 576

**Sum of ‘total tokens’:** 1624

## Experiment Timings

**Overall Experiment** started at 1766049991801, ended at 1766050004904, lasting 13103ms (13.10 seconds)

**Iteration #1** started at 1766049991802, ended at 1766049999076, lasting 7274ms (7.27 seconds)

**Iteration #2** started at 1766049999076, ended at 1766050004904, lasting 5828ms (5.83 seconds)