We will need a commitment scheme (see HWS). A (non-interactive) commitment scheme consists of two main algorithms (Commit, Verify):

- \( \text{Commit}(m, r) \rightarrow c \): Takes a message \( m \) and randomness \( r \) and outputs a commitment \( c \).
- \( \text{Verify}(m, c, r) \rightarrow b \): Checks if \( c \) is a valid opening to \( m \) (with respect to randomness \( r \)).

[The commitment scheme might also take public parameters (see HWS), but for simplicity, we omit them/leave them implicit.]

Requirements:
- Correctness: for all messages \( m \):
  \[ \Pr [ c \leftarrow \text{Commit}(m, r) : \text{Verify}(m, c, r) = 1 ] = 1 \]
  sampled uniformly

- Hiding: for all efficient adversaries \( A \), \( \Pr [ (m, c) \leftarrow A : m \neq m \text{ and } \text{Verify}(m, c, r) = 1 = \text{Verify}(m, c, r) ] = \text{negl} \).

- Binding: for all efficient adversaries \( A \), if \( \Pr [(m, m, c, r, c) \leftarrow A : m = m \text{ and } \text{Verify}(m, c, r) = 1 = \text{Verify}(m, c, r) ] = \text{negl} \).

\[
\rightarrow \text{We will require perfect binding [for every commitment } c, \text{ there is only 1 possible } m \text{ to which the power can open } c.\]

A ZK protocol for graph 3-coloring:

\[ \text{prover } (G) \quad \text{contains } n \text{ nodes, } m \text{ edges} \]

\[ \text{verifier } (G) \]

\[
\begin{align*}
\text{prover } (G) & \quad \text{let } K_i \in \{0,1,2\} \text{ be a 3-coloring of } G \\
& \quad \text{choose random permutation } \pi \text{ of } \text{rand}(\{0,1,2\}) \\
& \quad \text{for } i \in \pi: \\
& \quad \quad c_i \leftarrow \text{Commit}(K_i; r) \\
& \quad \text{for random } r: \\
& \quad \quad c_1, \ldots, c_n \\
& \quad \quad (K_i, r_i), (K_j, r_j) \\
& \quad (i, j) \in E \\
\text{verifier } (G) & \quad \text{accept if } K_i \neq K_j \text{ and } K_i, K_j \in \{0,1,2\} \\
& \quad \text{Verify}(K_i, c_i, r) = 1 = \text{Verify}(K_j, c_j, r) \\
& \quad \text{reject otherwise} \\
\end{align*}
\]

Intuitively: Prover commits to a coloring of the graph.
Verifier challenges prover to reveal coloring of a single edge
Prover reveals the coloring on the chosen edge and opens the entries in the commitment.

Completeness: By inspection, [if coloring is valid, prover can always answer the challenge correctly]

Soundness: Suppose \( G \) is not 3-colorable. Let \( K_1, \ldots, K_n \) be the coloring the prover committed to. If the commitment scheme is perfectly binding, \( c_1, \ldots, c_n \) uniquely determine \( K_1, \ldots, K_n \). Since \( G \) is not 3-colorable, there is an edge \((i, j) \in E\) where \( K_i = K_j \) or \( i \notin \{0,1,2\} \) or \( j \notin \{0,1,2\} \). Otherwise, \( G \) is 3-colorable with coloring \( K_1, \ldots, K_n \). Since the verifier chooses an edge to check at random, the verifier will choose \((i, j)\) with probability \( \frac{1}{|E|} \). Thus, if \( G \) is not 3-colorable,

\[ \Pr [\text{verifier rejects}] \geq \frac{1}{|E|} \]

Thus, this protocol provides soundness \( 1 - \frac{1}{|E|} \). We can repeat this protocol \( O(|E|^3) \) times sequentially to reduce soundness error to

\[ \Pr [\text{verifier accepts proof of false statement}] \leq \left( 1 - \frac{1}{|E|} \right)^{|E|^3} \leq e^{-|E|} = e^{-n} \quad \text{[since } 1+x \leq e^x \]
Zero knowledge: We need to construct a simulator that outputs a valid transcript given only the graph $G$ as input.

Let $V^*$ be a (possibly malicious) prover. Construct simulator $S$ as follows:

1. Choose $K_i \leftarrow \{0,1,2,3\}$ for all $i \in [n]$.
   Let $c_i \leftarrow \text{Commit}(K_i; r_i)$
   Given $(c_1, \ldots, c_n) \to V^*$.

2. $V^*$ outputs an edge $(c_i) \in E$
3. If $K_i \neq K_j$, then $S$ outputs $(K_i, K_j, r_i, r_j)$.
   Otherwise, restart and try again (if fails $k$ times, then abort)

Simulator succeeds with probability $\frac{2}{3}$ (over choice of $K_1, \ldots, K_n$). Thus, simulator produces a valid transcript with prob. $1 - \frac{1}{3^2} = 1 - \frac{1}{27}$ after $k$ attempts. It suffices to show that simulated transcript is indistinguishable from a real transcript.

   - Real scheme: prover opens $K_i, K_j$ where $K_i, K_j \leftarrow \{0,1,2,3\}$ [since prover randomly permutes the colors]
   - Simulation: $K_i$ and $K_j$ sampled uniformly from $\{0,1,2,3\}$ and conditioned on $K_i \neq K_j$, distributions are identical

In addition, $(c_i)$ output by $V^*$ in the simulation is distributed correctly since commitment scheme is computationally-hiding (e.g. $V^*$ behaves essentially the same given commitments to a random coloring as it does given commitment to a valid coloring).

If we repeat this protocol (for soundness amplification), simulator simulate one transcript at a time.

**Summary:** Every language in NP has a zero-knowledge proof.

In many cases, we want a stronger property: the prover actually "knows" why a statement is true (e.g., it knows a "witness").

For instance, consider the following language:

$$L = \{ h \in G \mid \exists x \in \mathbb{Z}_p : h = g^x \} \in G$$

Note: this definition of $L$ implicitly defines an NP relation $R$:

$$R(h, x) = 1 \iff h = g^x \in G$$

In this case, all statements in $G$ are true (i.e., contained in $L$), but we can still consider a notion of proving knowledge of the discrete log of an element $h \in G$ — conceptually stronger property than proof of membership.

**Philosophical question:** What does it mean to "know" something?

If a prover is able to convince an honest verifier that it "knows" something, then it should be possible to extract that quantity from the prover.

**Definition.** An interactive proof system $(P, V)$ is a proof of knowledge for an NP relation $R$ if there exists an efficient extractor $E$ such that for any $x$ and any prover $P^*$

$$\Pr \left[ \omega \leftarrow E^{P^*}(x) : R(x, \omega) = 1 \right] \geq \Pr \left[ \langle P^*, V \rangle(x) = 1 \right] - \epsilon$$

*more generally, could be polynomially smaller*