Theorem 1. *Subset-sum is $\mathcal{NP}$-complete.*

Proof. You proved in the assignment that 1-in-3-SAT is $\mathcal{NP}$-complete.[1] It's easy to see that subset-sum is in $\mathcal{NP}$, and we show a Karp reduction from 1-in-3-SAT.

Given a 3CNF formula Φ over variables $x_1, ..., x_n$ with m clauses $C_1, ..., C_m$, we construct a set of numbers $w_1, ..., w_{2n}$ and a target number W such that Φ is satisfiable if and only if there is $S \subseteq [2n]$ such that $\sum_{i \in S} w_i = W$.

> **High-level idea:**
>
> Intuitively, we think of "$w_i \in S$" as setting the i^{th} variable to true and of "$w_{n+i} \in S$" as setting the i^{th} variable to false. We'd like a sum $\sum_{s \in S} w_s = W$ to represent an assignment to the variables.
>
> We will build W that enforces the condition that the sum represents a consistent assignment (i.e., either w_i or w_{n+i} is chosen for each $i \in [n]$) and that for each clause, exactly one literal in the clause is true.
>
> **Technical setup:**
>
> We think of all numbers in the proof in base $B = 4$ (we'll see later why we chose 4), and consisting of $n + m$ digits. That is, we represent each number as $w_i = \sum_{j \in [n+m]} a_j \cdot B^j$ for some coefficient a_j's in $\{0, 1, 2, 3\}$.
>
> The first n digits will represent variables, and the last m digits will represent clauses. The w_i's will have all digits set to either zero or one, and summing w_i's is equivalent to summing their digits.
>
> When summing w_i's and looking at the result, we'll interpret a digit $i \in [n]$ being set to 1 as "the variable has been assigned" and a digit $(n + k) \in \{n + 1, ..., n + m\}$ being set to 1 as "the clause has been satisfied".

The numbers and target-sum. For each $i \in [n]$, each of w_i, w_{n+i} has:

- Its i^{th} digit set to 1.
- For every clause C_k that includes x_i, the $(n + k)^{th}$ digit of w_i is set to 1.
- For every clause C_k that includes $\neg x_i$, the $(n + k)^{th}$ digit of w_{n+i} is set to 1.

The target W has all its digits set to 1 (i.e., $W = \sum_{j \in [n+m]} B^j$).

For runtime, convince yourself that given a formula Φ, we can output the w_i's and W in time $\text{poly}(n, m)$.

[1] Recall that in 1-in-3-SAT the input is a 3CNF formula Φ, and we need to decide if there is x such that for every OR in Φ *exactly one* of the three literals in the OR is set to true by x.

Analysis. If there's $x_1, ..., x_n$ satisfying Φ, we construct S such that $\sum_{s \in S} w_s = W$. Specifically, for $i \in [n]$, if $x_i = 1$ we add w_i to S and otherwise we add w_{n+i} to S.

Note that $\sum_{s \in S} w_s$ has, in each of the first n positions, exactly one digit set to 1 (because for every $i \in [n]$ either w_i or w_{n+i} has been added to S). Also, since each clause C_k is satisfied by *exactly one* literal in the clause, there will be exactly one w_i or w_{n+i} in S whose $(n+k)^{th}$ digit is set to 1. Hence, each of the last m digits of $\sum_{s \in S} w_s$ is set to 1, so overall $\sum_{s \in S} w_s = \sum_{j \in [n+m]} B^j = W$.

In the reverse direction, let S such that $\sum_{s \in S} w_s = W = \sum_{j \in [n+m]} B^j$. We claim that each 1-digit of W came from a single $w_s \in S$. The only reason why this might not happen is wraparound (i.e., if $B^{j+1} = c_j \cdot B^j$), and to see that this is impossible, recall that for every digit $1, ..., n+m$ there are at most three numbers with 1 in that digit,[2] so the sum $\sum_{s \in S} w_S$ can contribute at most 3 to each digit of W, and our choice of $B > 3$ prevents wraparound (indeed, this is the reason we chose $B = 4$).

So now we're certain that for every position $1, ..., n+m$ there is exactly one $w_s \in S$ with 1 in this position. In particular, for every $i \in [n]$ either w_i or w_{n+i} is in S, and we assign $x_1, ..., x_n$ accordingly (i.e., $x_i = 1$ if and only if $w_i \in S$). And for each position $n+k$ has 1 there is a single w_i or w_{n+i} with 1 in that position, exactly one literal in each clause is satisfied by this assignment, hence Φ is satisfied by $x_1, ..., x_n$. ■

[2] For every $i \in [n]$ there are only w_i and w_{n+i} with 1 in this position, and for every $k \in [m]$ there are only three literals participating in C_k, hence only three w_i's and/or w_{n+i}'s with 1 in position $n+k$.