Subset Sum is NP-complete

$SUBSET\text{-}SUM = \{(S, t) \mid S$ has a subset whose elements add up to $t\}$.

Here S is a set of integers and $t \in \mathbb{Z}$

We already showed $SUBSET\text{-}SUM$ is in NP.

Vertex Cover Reduces to Subset Sum

(Note: The book uses 3-SAT)

Given an instance (G, k) of VC we map it to an instance (S, t) of $SUBSET\text{-}SUM$.

Suppose $G = (V, E)$. We assume that G is represented in "incidence matrix" form, i.e., there is a matrix A with a row for each node, and a column for each edge Entry A_{ij} is 1 if node i is one of the endpoints of edge j, and 0 otherwise.

Example

$G = (\{A, B, C, D, E\}, (A, B), (A, C), (A, D), (B, C), (B, E), (C, E), (D, E)\}$
$k = 3$

	(A,B)	(A,C)	(A,D)	(B,C)	(B,E)	(C,E)	(D,E)
A	1	1	1	0	0	0	0
B	1	0	0	1	1	0	0
C	0	1	0	1	0	1	0
D	0	0	1	0	0	0	1
E	0	0	0	0	1	1	1

The Mapping – Part I

We first map the matrix A to a matrix A' as follows:

1. Put 1 in front of each row.

2. For each column j, add a row with all 0's and one 1 in the j^{th} position (we call this an "edge row")

3. Put an all-0 column in between each of the columns

The Matrix A'

		(A,B)	(A,C)	(A,D)	(B,C)	(B,E)	(C,E)	(D,E)
A	1	0 1	0 1	0 1	0 0	0 0	0 0	0 0
B	1	0 1	0 0	0 0	0 1	0 1	0 0	0 0
C	1	0 0	0 1	0 0	0 1	0 0	0 1	0 0
D	1	0 0	0 0	0 1	0 0	0 0	0 0	0 1
E	1	0 0	0 0	0 0	0 0	0 1	0 1	0 1
		0 1	0 0	0 0	0 0	0 0	0 0	0 0
		0 0	0 1	0 0	0 0	0 0	0 0	0 0
		0 0	0 0	0 1	0 0	0 0	0 0	0 0
		0 0	0 0	0 0	0 1	0 0	0 0	0 0
		0 0	0 0	0 0	0 0	0 1	0 0	0 0
		0 0	0 0	0 0	0 0	0 0	0 1	0 0
		0 0	0 0	0 0	0 0	0 0	0 0	0 1

The Mapping – Part II

1. Interpret each row of A' as a binary number to give S.
2. Let t be the binary representation of k followed by a string of $|E|$ repetitions of 10.

So for $k = 3$, $t = 11\ 10\ 10\ 10\ 10\ 10\ 10\ 10$

The Mapping is Poly-time

For an input $w = ((V, E), k)$, $f(w)$ has $|V| + |E| + 1$ numbers of length $2|E| + 1$ each. It is not hard to see that the mapping is poly-time in $|V| + |E|$.

If $w \in VC$, Then $f(w) \in SUBSET\text{-}SUM$

Proof: If there is a vertex cover of size k, we show there is a set $S' \subseteq S$ which sums to t. There is a set of nodes of size k which "hits" every edge at least once. For every node in the cover, put the number corresponding to the row for that node into S'. Also, if some edge is incident to *only one* node in the cover, put the corresponding "edge row" into S'. It is not hard to see that the elements of S' add up to t: for each edge, the will be exactly two 01 pairs in the corresponding column – if the edge is incident to two nodes in the cover, the 01's will come from the corresponding node rows. Otherwise one will come from the node row for the node which covers it, and the other will come from the corresponding edge row. These add up to give a 10 pair in t

Example Using Cover A, C, E

	(A,B)	(A,C)	(A,D)	(B,C)	(B,E)	(C,E)	(D,E)
A 1	0 1	0 1	0 1	0 0	0 0	0 0	0 0
B 1	0 1	0 0	0 0	0 1	0 1	0 0	0 0
C 1	0 0	0 1	0 0	0 1	0 0	0 1	0 0
D 1	0 0	0 0	0 1	0 0	0 0	0 0	0 1
E 1	0 0	0 0	0 0	0 0	0 1	0 1	0 1
	0 1	0 0	0 0	0 0	0 0	0 0	0 0
	0 0	0 1	0 0	0 0	0 0	0 0	0 0
	0 0	0 0	0 1	0 0	0 0	0 0	0 0
	0 0	0 0	0 0	0 1	0 0	0 0	0 0
	0 0	0 0	0 0	0 0	0 1	0 0	0 0
	0 0	0 0	0 0	0 0	0 0	0 1	0 0
	0 0	0 0	0 0	0 0	0 0	0 0	0 1

If $f(w) \in SUBSET\text{-}SUM$ Then $w \in VC$

Proof: If there is an S' which sums to t, note that there can't be any carries from any column except the first, since there are no more than 3 1's in any column of S and there's a column of 0's between each nonempty column. So exactly k node rows must be selected. Make these the nodes in the cover. Since the numbers in S' sum to t, there must be exactly two 01's corresponding to each edge, and at least one of which is a from a node row, so each edge must be covered.

PARTITION is NP-complete

$PARTITION =$
$\{U \mid U$ contains a subset which sums to half the sum of all elements$\}$.

HINT: Use SUBSET-SUM
There are three cases:
$t = 1/2$ the sum of integers in S.
$t < 1/2$ the sum of integers in S.
$t > 1/2$ the sum of integers in S.