



OECD International Academy for Tax Crime Investigation

*Anti-Money Laundering: Current Trends, Prosecutions,
and the Challenges around Cryptocurrencies*



Hierarchical Deterministic Wallets and Seeds

BIP-32

HIERARCHICAL DETERMINISTIC WALLETS

Hierarchical Deterministic Wallets

- Hierarchical Deterministic Wallets are wallets that can be easily backed up with seed word lists.
- If you lose the device holding your wallet, you can recreate the wallet using your seed words, and regain access to your funds.
- You can use the seed words to create the same wallet on several different devices.
- **Most modern wallets are hierarchical deterministic wallets.**

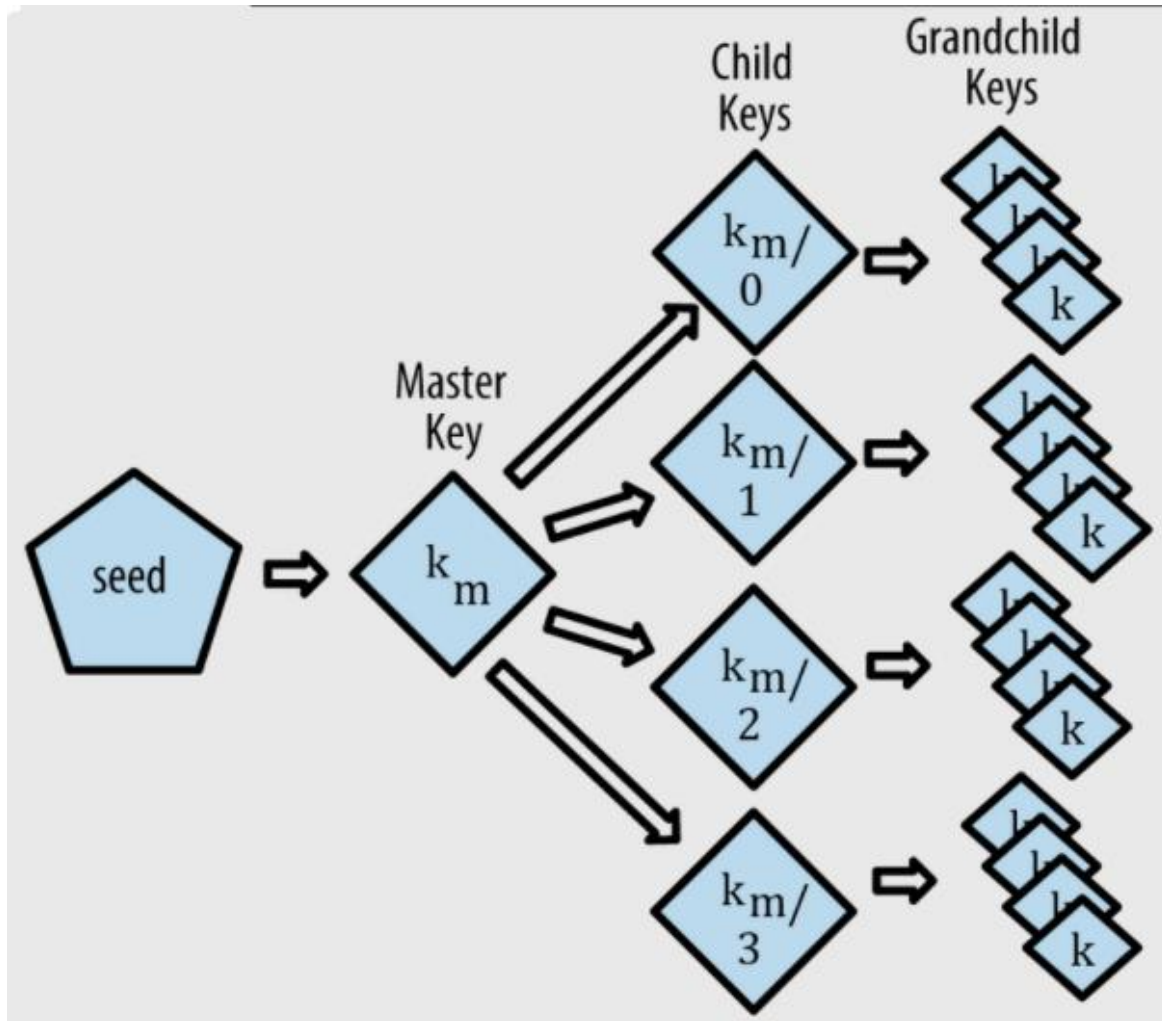
Hierarchical Deterministic Wallets (HD)

- **Hierarchical:** because all keys and addresses in the wallet are derived from one initial “**master seed**” (i.e. a very large number).
 - **The master seed is used to generate the initial key**
 - **Sub-keys are generated from the initial key** (i.e. child keys, grandchild keys, etc.)
- **Deterministic:** if you start with the same initial seed when you install a wallet, the master key and subkeys which will be derived from it will be identical.

Master Seed = Number = Seed Words List

- **Master seeds are a random number**
- **The number is represented by a list of words.**
 - It is easier for humans to remember a list of words than large numbers.
- **Wallets typically use 12 or 24 words to represent a master seed.**
 - **12 words** are primarily used for cell phone and computer wallets.
 - **24 words** are primarily used for hardware wallets.
- Some wallets allow you to use a pass phrase as well as seed words in order to generate the master seed.

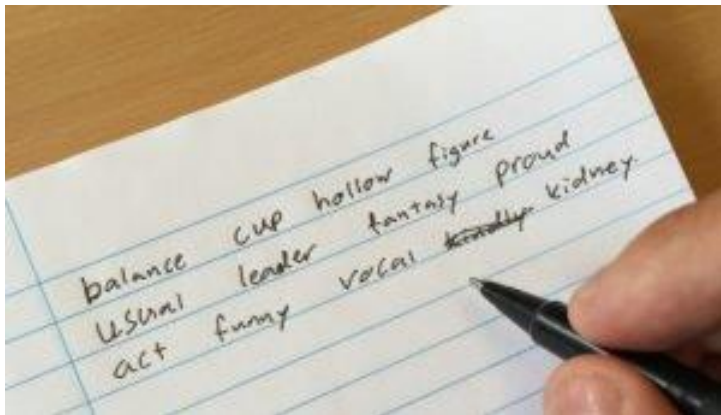
Seeds and Keys



SEED WORD LISTS

Seed Word Lists

- Lists of words which function as wallet backups.



Wallets and Seeds

- Seed word lists are used to back up most digital wallets:

Cell phone wallets

Computer wallets

Hardware wallets

Browser plugin/extension wallets

- Seed word lists aren't used to back up some wallets:

Paper wallets

Custodial wallets

When you set up a wallet

- **Your wallet software will supply you with a word list if you are setting up a new wallet.**
 - Wallets do not come with seed word lists coded into them.
 - Wallets use a number of “environmental factors” to generate a random number.
 - The random number determines which words to give you.
- Both hardware wallets and software wallets supply you with a seed word list when you set them up.

How seed word lists are generated

(short version)

- During the wallet installation a random number is generated.
- Some mathematics are applied to the number.
- The resulting number is split into 12 smaller numbers.
- Each smaller number maps to one of 2048 specific words.

How seed word lists are generated

(slightly longer explanation...)

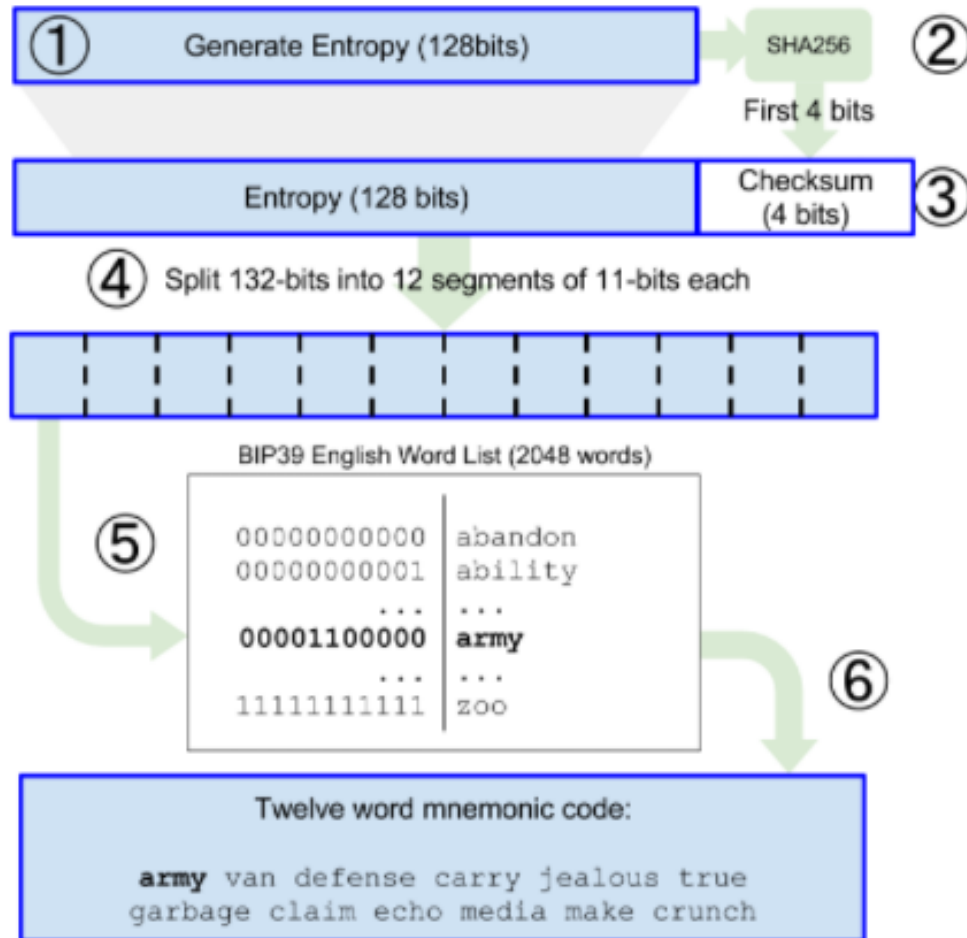
1. During installation the wallet generates a random 128-bit number (i.e. the Seed).
2. The random number is hashed (SHA-256)
3. The first 4 bits of the hash are added to the end of the random number.
4. The resulting 132-bit number is split into twelve 11-bit numbers.
5. Each word in the BIP-39 word list maps to a specific number.
6. The ordered sequence of words represents the initial seed number.

12 Seed Words = 132 Bit Number

1 1 0 0 1 0 0 0 1 0 0 0 1 0 0 0 1 1 0 1 1 1 0 1 0 1 0 0 0 1 1 0 0 0 1 1 0 0 0 1 0 1 1 0
1 0 0 0 0 0 1 1 0 1 0 0 1 1 1 1 1 0 0 1 1 0 1 1 1 0 1 0 1 1 0 0 0 0 1 1 1 0 0 1 1 0 0
0 0 0 1 1 0 1 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 1 0 0 0 1 1 1 1 0 1 0 0 1 1 1 1 1 0 1 0 1

Each 11 bits are represented by a different word

Mnemonic Words 128-bit entropy/12-word example



Antonopoulos, A. and G. Wood (2018). Mastering Ethereum. Implementing Smart Contracts, O'Reilly.

Seeds Can Generate Keys for Different Coin Types

Registered coin types

These are the registered coin types for usage in level 2 of BIP44 described in chapter "Coin type".

All these constants are used as hardened derivation.

index	hexa	symbol	coin
0	0x80000000	BTC	Bitcoin
1	0x80000001		Testnet (all coins)
2	0x80000002	LTC	Litecoin
3	0x80000003	DOGE	Dogecoin
4	0x80000004	RDD	Reddcoin
5	0x80000005	DASH	Dash (ex Darkcoin)
6	0x80000006	PPC	Peercoin

BIP-39

SEED WORD LIST STANDARD

BIP-39 Standard

- Changes made to Bitcoin are made and adopted via proposals called Bitcoin Improvement Proposals (BIP).
- BIP-39 is a proposed and accepted standard for creating wallet backups using seed word lists.
- BIP-39 includes lists of seed words in 10 different languages
- Most wallets follow the BIP-39 standard, but some do not.

Wordlists (BIP-39)

Ten Languages

2048 words in each language

- English
- Japanese
- Korean
- Spanish
- Portuguese
- Chinese (Simplified)
- Chinese (Traditional)
- French
- Italian
- Czech

1	abandon		
2	ability	2041	yellow
3	able	2042	you
4	about	2043	young
5	above	2044	youth
6	absent	2045	zebra
7	absorb	2046	zero
8	abstract	2047	zone
9	absurd	2048	zoo

1	的	2040	嘗
2	一	2041	卿
3	是	2042	妨
4	在	2043	艇
5	不	2044	吞
6	了	2045	韋
7	有	2046	怨
8	和	2047	矮
9	人	2048	歇

2048 lines (2048 sloc)	2048 lines (2048 sloc)	2048 lines (2048 sloc)	2048 lines
1 abandon	1 abaisser	1 abaco	1 的
2 ability	2 abandon	2 abbaglio	2 一
3 able	3 abdiquer	3 abbinato	3 是
4 about	4 abeille	4 abete	4 在
5 above	5 abolir	5 abisso	5 不
6 absent	6 aborder	6 abolire	6 了
7 absorb	7 aboutir	7 abrasivo	7 有
8 abstract	8 aboyer	8 abrogato	8 和
9 absurd	9 abrasif	9 accadere	9 人
10 abuse	10 abreuver	10 accenno	10 这
11 access	11 abriter	11 accusato	11 中
12 accident	12 abroger	12 acetone	12 大
13 account	13 abrupt	13 achille	13 为
14 accuse	14 absence	14 acido	14 上
15 achieve	15 absolu	15 acqua	15 个
16 acid	16 absurde	16 acre	16 国

<https://github.com/bitcoin/bips/tree/master/bip-0039>

Some Wordlist Considerations (BIP-39)

- **Smart selection of words**
 - The wordlist is created in such way that it's enough to type the ***first four letters*** to unambiguously identify the word
- **Similar words avoided**
 - Word pairs like "build" and "built", "woman" and "women", or "quick" and "quickly" not only make remembering the sentence difficult, but are also more error prone and more difficult to guess

Note:

Some wallets do not use BIP-39

- **Electrum Wallet** does not use BIP-39
 - Created 2 years before BIP-39 standard proposed
 - Electrum derives its private keys and addresses from a seed phrase made of natural language words.
 - Electrum uses a different seed derivation algorithm.
 - New installs of Electrum now use the standard BIP-39 word lists.

DERIVATION PATHS

Derivation Paths

- There are different paths which can be used to derive keys from the master key.
- The default derivation path for Bitcoin is `m / 44' / 0' / 0' / 0`.
- Each number in that path represents a certain level and path in the tree.

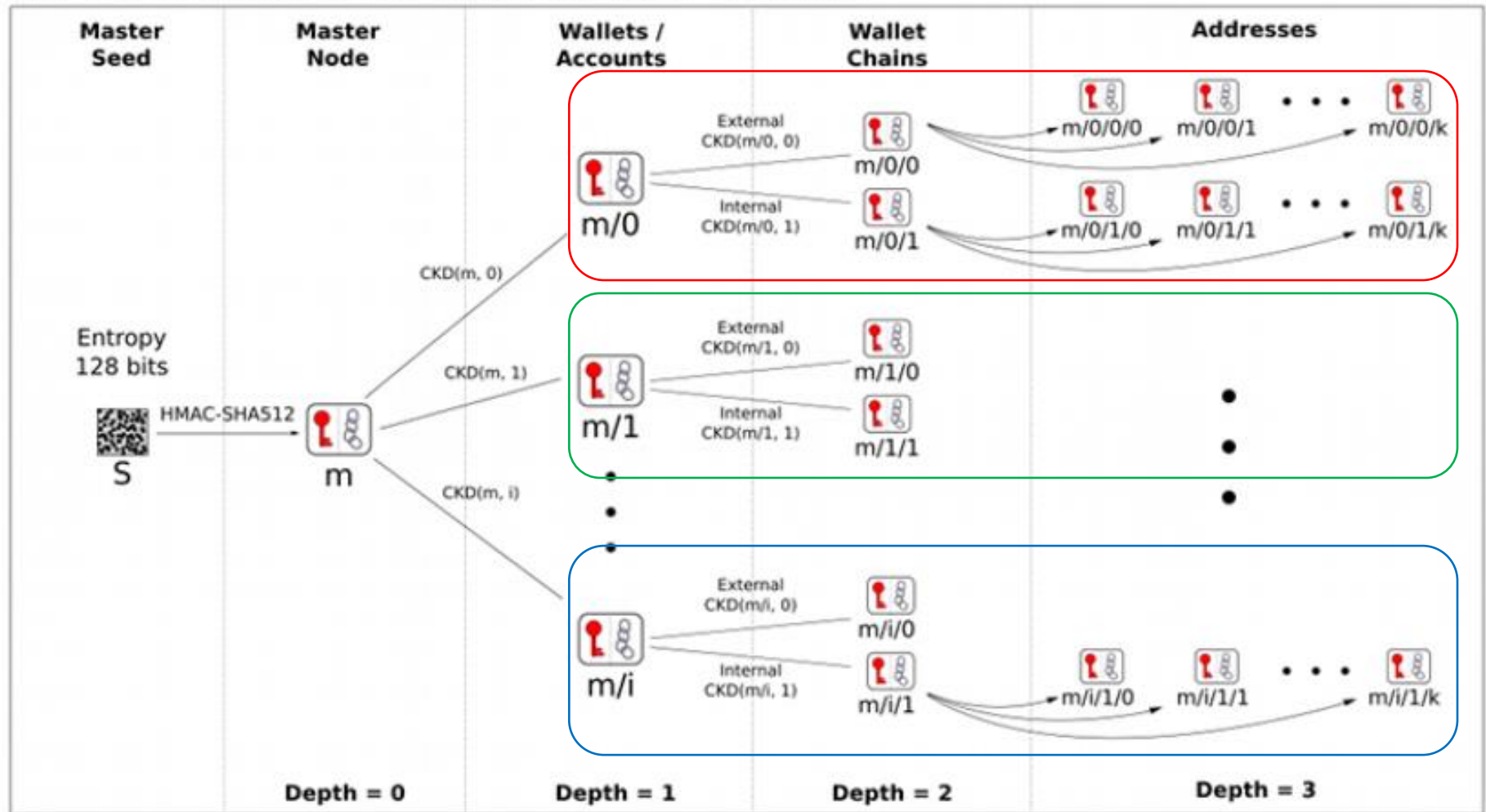
`m / purpose' / coin_type' / account' / chain / address_index`

Different wallets use different derivation paths.

When you use a seed word list to re-create a wallet, you need to use the right wallet.

Derivation Paths

BIP 32 - Hierarchical Deterministic Wallets



Child Key Derivation Function $\sim CKD(x,n) = HMAC-SHA512(x_{Chain}, x_{PubKey} || n)$

What Seed Word Lists and Derivation Paths Mean for an Investigator

- **If you know a person's seed word list:**
 1. You can generate all their keys and addresses.
 2. You can control spend their cryptocurrency
 3. You can map out all their transactions
- **You need to know their wallet type:**
 - When you re-create the suspect's wallet, you need to use a wallet which uses the same **derivation paths**.
- **You don't need to know their password**
 - Their password is used to encrypt their wallet after it has been created.

Resources – YouTube Videos

(Andreas Antonopoulos)

- [Bitcoin Q&A: How do mnemonic seeds work?](#)
- [Bitcoin Q&A: Passphrases and seed storage](#)

Bitcoin Improvement Proposals (BIP)

- **BIP-32:** Hierarchical Deterministic Wallets
- **BIP-39:** Mnemonic code for generating deterministic keys
- **BIP-44:** Multi-Account Hierarchy for Deterministic Wallets
- **BIP-49:** Derivation scheme for P2WPKH-nested-in-P2SH based accounts