DOI:10.1145/2701411

The myths, the hype, and the true worth of bitcoins.

BY AVIV ZOHAR

Bitcoin: Under the Hood

"I JUST WANT to report that I successfully traded 10,000 bitcoins for pizza," wrote user laszlo on the Bitcoin forums in May 2010—reporting on what has been recognized as the first item in history to be purchased with bitcoins.[a] By the end of 2013, about five years after its initial launch, Bitcoin has exceeded everyone's expectations as its value rose beyond the $1,000 mark, making laszlo's spent bitcoins worth millions of dollars. This meteoric rise in value has fueled many stories in the popular press and has turned a group of early enthusiasts into millionaires.

Stories of Bitcoin's mysterious creator, Satoshi Nakamoto, and of illegal markets hidden in the darknet have added to the hype. But what is Bitcoin's innovation? Is the buzz surrounding the new cryptocurrency justified, or will it turn out to be a modern tulip mania? To truly evaluate Bitcoin's novelty, its potential impact, and the challenges it faces, we must look past the hype and delve deeper into the details of the protocol.

Bitcoin, a peer-to-peer digital cryptocurrency launched in 2009, has been slowly growing. Nakamoto described the protocol in a white paper published in late 2008[25] and released the software as an open source project, which has since been maintained by a large number of developers, most of them volunteers. Bitcoin's network and its surrounding ecosystem have grown quite substantially since its initial release. Its dollar value, which most will admit is largely based on speculation on its future worth, has been extremely volatile. The currency had gone through several hype-driven bubbles and subsequent devaluations, attaining higher values each time.

Bitcoin's promise is mainly a result of the combination of features it bundles together: It is a purely digital currency allowing payments to be sent almost instantly over the Internet with extremely low fees. Like cash, it is nearly anonymous, and transactions are effectively irreversible once they are committed. Bitcoin addresses (the

» key insights

- Bitcoin's operation relies on the Block Chain—a distributed ledger of transactions that is synchronized between all nodes. The main challenge the protocol successfully tackles is to ensure nodes agree on the contents of this ledger.
- Going forward, the protocol faces challenges in several domains: ensuring the privacy of users, scaling up to high throughput, maintaining mining decentralization, more easily deploying updates to the core protocol, increasing the robustness of its overlay network, and structuring rewards within the protocol to improve incentives.
- Continuous innovations are slowly addressing these challenges. Along with applications outside of the economic domain, Bitcoin may yet fulfill its promise to become a meaningful force in the global money transmission market.

[a] https://bitcointalk.org/index.php?topic=137.0

IMAGE BY ASAF ELIASON

equivalent of accounts) are free, and anyone can open as many as they would like. Set apart from other existing forms of digital currency, Bitcoin is based on a decentralized protocol: There is no organization or government in control of its operation. As a consequence, there is no central entity able to apply monetary policy, and its supply has been set in advance—there will never be more than 21 million bitcoins.

Without the initial support of a government or some other large central entity, initial adoption has been slow. Early adopters experienced the negative side of the network effect: having relatively few places to spend bitcoins, or to acquire them has made them less useful. The uncertain regulatory and legal status, the failure of many exchanges,[12,23] as well as the initial lack of user-friendly software wallets have also hindered growth.

All this is slowly changing. Exchanges that trade local currencies for bitcoins have appeared in more places, including ATMs that exchange bitcoins for cash. Digital wallets with improved interfaces can be found in app stores, and point-of-sale systems now allow any business to accept bitcoins more easily than ever before. Progress has also been made on legal and regulatory aspects. In some countries it is now clear how bitcoin transactions are taxed, and regulators have started to draft guidelines for exchanges and banks (most notably, New York's so-called BitLicense[10] has been recently put into effect). From a security standpoint, Bitcoin's core protocol and its network have been surprisingly resilient and have not been successfully compromised to date, adding to the confidence in its foundations.[b]

Will Bitcoin expand to become a substantial part of the payments market, or will it vanish as a passing trend?

[b] Other systems that use bitcoins have been hacked, and large sums of money have been stolen.

Only time will tell. While not without its flaws, Bitcoin does not need to be perfect to become prevalent—no system is—it need only compete with the alternatives; cash, credit cards, and wire transfers all have their downsides and imperfections. But whether or not it survives, Bitcoin's grand experiment promises to have a deep impact on the way we think of financial systems.

Bitcoin's core innovation, which may yet extend its impact beyond digital currencies, lies at the heart of a well-known problem in computer science, namely, the Byzantine consensus problem. Dealing with the synchronization of information in a distributed system in the presence of malicious adversaries, Byzantine consensus[17] has been extensively researched and several algorithms and impossibility results are known. Bitcoin's main contribution amounts to a solution to a variant of the consensus problem—one that does not require participants to have strong identities, but instead relies on assumptions that limit the computational resources of attackers.[22] But what does agreement over information in a distributed system have to do with money? To explain, we must first discuss the traditional design of digital money, and how Bitcoin's design differs.

Digital money, double spending, and the intermediary. Any viable medium of exchange must have a limited supply. Physical forms of money have always had this property. Precious metals, much to the dismay of alchemists, could not be easily produced, and modern bank notes have had many anti-forgery countermeasures embedded in them. In the age of the Internet, digital money has a clear advantage: it is faster to transmit. Unfortunately, information cannot easily replace physical tokens. It is too easy to replicate, and anyone who uses some string of bits as payment would be able to keep a copy, and perhaps use it to pay someone else. This problem, which is inherent to digital currencies, is known as the double spending problem.[c]

The classic solution to double spending, one at the foundation of most modern online banking systems, is to do away with tokens altogether. Money is, after all, a form of memory representing who has provided services and goods to others. Instead of holding physical tokens that represent credit, it is possible to list the holdings of each individual in a ledger. Transferring money is then accomplished simply by changing the records on the ledger—stored in the memory of some server—adding to one account balance, and subtracting from the other.

This design adds a third party to all transactions—the record keeper. This intermediary is given a great deal of power: It can refuse to carry out certain transfers, to change balances even without the consent of the transacting parties, or to demand high fees in exchange for its indispensable services, something that had never been possible with physical forms of money. Additionally, in contrast to the anonymity of cash transactions, the privacy of individuals transacting with digital currency is compromised. The intermediary itself is explicitly notified of every transaction that takes place. Finally, the existence of record keepers through which all payments are funneled allows for government intervention and regulation. Regulation, which serves to hinder criminal activity and to guard against misuse of the funds by the intermediary itself, has its downsides. Regulatory compliance imposes a direct cost on organizations, and also introduces barriers that restrict entry to the money transmission market, reduce competition, and so serve to increase fees even further.

Bitcoin seeks the best of both worlds: to enjoy the full benefits of the digital domain, but also to greatly weaken any third party through competition and decentralization. Most of Bitcoin's features are natural implications of this choice: the inability to reverse payments and the fixed supply of money, for example, are natural design choices when there is no centralized entity that can verify money has been stolen and payments should be reversed, or whether or not more money should be issued. Many other beneficial properties of Bitcoin are achieved by the application of more modern practices: Unlike credit cards that bear the burden of backward compatibility and have card numbers and expiration dates that are easy to steal,[2] access to bitcoins is guarded by public key cryptography. Other advantages of Bitcoin are due to its open nature. The open source model boosts its transparency, adds confidence in its stability and security, and grants open access to its APIs, which enable agile development within the surrounding ecosystem.

Replacing the intermediary. In order to reduce the influence of any third party but still allow funds to be transferred, Bitcoin's design replaces the centralized intermediary with many weaker entities that maintain the ledger. The nodes in charge of Bitcoin's transaction processing, also known as miners, form a large and connected peer-to-peer network that together authorizes all money transfers. Each miner checks the actions of others to ensure money is not mishandled, and competes to authorize a share of the transactions.

One of the main goals of the protocol's design is to make it easy to join the network. Anyone can download the open source software and use readily available hardware to run a node. Nodes connect to each other via TCP connections over the Internet and share the IP addresses of other known peers to form a robust distribution network. Thus, no one is granted absolute control of the system, and no single entity is able to block transactions, or to extract unreasonably high fees.

Unlike a distributed design that aims to share the load among many machines, Bitcoin nodes do not partition the workload or the ledger among them. In fact, in order to allow each node to act independently, data is massively replicated, and each participant repeats all verification work. This replication naturally requires all nodes to be notified of every transaction, as each transfer of funds must be recorded in all copies of the ledger. Users who wish to send money create a message requesting the transfer. Transfers are made between Bitcoin addresses, which act as the approximate equivalent of an "account" (each address is in fact the hash of a cryptographic public key). The sender's message is digitally signed to prove ownership of the funds, and is transmitted to some of the nodes in the network. Nodes verify the signatures and then forward the message to their peers, ensuring it is sent to the entire network. As

[c] The exact state of a quantum bit cannot be copied, and so quantum currency systems that disallow double spending are theoretically possible.[1,31]

a consequence, all bitcoin transactions are public.

Notice that nothing stops the owner of funds from creating and signing two conflicting transaction messages that transfer the same funds to different destination addresses. This, in fact, is the double spending problem as it manifests itself in Bitcoin. Nodes that have received these transactions may adopt different ones and consequently disagree about the state of the ledger. This is where Bitcoin's main innovation is rooted, at the synchronization of information in its ledger.

The Block Chain

Bitcoin's main data structure, the block chain, is the key to understanding how information consistency is maintained between nodes and how conflicts are resolved. The block chain, as its name suggests, is composed of blocks—batches of approved transactions that have been grouped together. Each block contains the cryptographic hash of its predecessor that, for all intents and purposes, serves as a unique identifier of the previous block (hash collisions are very rare, and difficult to find—an important property of cryptographic hash functions).

The block chain thus forms an incremental log of all transactions that have ever occurred since the creation of Bitcoin, starting with the "Genesis Block"—the first block in the chain. If one reads the log from start to finish, every transfer of money can be verified and funds can be followed to compute the balance of each Bitcoin address. Nodes that were offline can easily catch up by requesting only the few recent blocks that they are missing.

The block chain grows steadily as new blocks extend it, referencing their predecessors, and adding new transactions. Newly created blocks are flooded within the network to ensure all nodes possess them. In order to maintain the consistency of the chain, valid blocks are only allowed to include transactions consistent with current balances as determined by their predecessors in the chain.

If all nodes possess the exact same copy of the block chain, then all is well; the ownership of every fraction of a bitcoin is known and agreed upon by everyone. This, unfortunately, is not always the case. The network is distributed, and the creation of blocks is uncoordinated. Thus, blocks that are formed approximately at the same time by different nodes may extend the same parent block and create a fork in the chain. Such blocks represent a different version of the transaction log, and are likely to contain conflicting transactions (for example, see Figure 1).

Bitcoin's promise is mainly a result of the combination of features it bundles together: It is a purely digital currency allowing payments to be sent almost instantly over the Internet with extremely low fees.

We have finally reached the core of the Bitcoin protocol: its mechanism for selecting between conflicting histories. The mechanism consists of two main rules that govern block creation and adoption:

1. *Block creation is difficult (by design).* Valid blocks are required to contain a proof-of-work: the solution to a computationally hard problem. A solution to the problem is easily verifiable, but finding it requires many guesses to be made, and takes a long time (the problem itself is based on the repeated application of cryptographic hashing to the block's contents; see the sidebar "Bitcoin's Proof-of-Work" for additional information).

2. *Adopt the longest chain.* Blocks are distributed throughout the network. When nodes learn about conflicting blocks that make up a longer consistent chain, they adopt it and abandon blocks in their shorter version.[d]

The two rules work together to bring the network to consensus regarding the history of transactions. First, as blocks are rarely created, few conflicts occur to begin with: If block creation is infrequent, a new block will most likely be propagated to all nodes before the next one is produced. The next block will thus reference it and will not include conflicting transactions. The difficulty of the computational problem is automatically adjusted so blocks are created only once every 10 minutes in expectation throughout the entire network. This period of time is sufficiently long to make conflicts extremely rare.

The longest-chain rule resolves these conflicts and ensures the net-

[d] In fact, nodes do not pick the longest chain, but rather the chain that contains the highest aggregate difficulty of proof-of-work computations (this measure is used because the difficulty of block creation is regularly adjusted by the protocol). It is simpler, however, to think instead of the length of the chain as this aggregate measure.

Bitcoin's Proof-of-Work

To make block creation difficult, the protocol requires the cryptographic hash of each block (or, to be more precise, the hash of the block's header) will be a small number under some threshold (called "the target"). The block's header contains a field called the nonce that can contain an arbitrary string of bits. If the block's hash is too large, the nonce can be changed and the hash can be recomputed. An important property of strong cryptographic hash functions is that a change to even a single bit in their input completely and unpredictably changes their output. Many attempts are thus needed to find a nonce that will fit the block, and produce a hash below the target. For example, if the target has 60 zeros in its most significant bits, fewer than one in 2^{60} attempted hashes will result in a successful attempt, requiring miners to perform a great deal of computational operations to create a valid block.

In addition to changes to the nonce, every change to the contents of the block also changes its hash, so once a match is found, the block cannot be modified. The proof-of-work can be easily verified by each node that later receives the block, simply by checking its hash value.

To ensure blocks are created in expectation once every 10 minutes, the threshold for successful block creation is adjusted automatically every 2,016 blocks. If, for example, blocks were created too quickly (as is often the case if additional hashing power was added to the network since the previous adjustment) the difficulty is raised by lowering the target threshold.

work will eventually converge to a single choice: If two conflicting blocks exist, each node in the network adopts one of them, but not the other, as the alternative is not currently part of a longer chain. The network is thus partitioned to nodes that accept one version of events, or the other. Once another block is created by one of the nodes, the tie is broken, and one of the possible versions of transaction history becomes longer. This version will then propagate and be adopted by the entire network. Ties among conflicting chains may in fact last longer, but eventually, due to the random nature of the computation involved in the block creation process, one chain will win the race, and the other will be abandoned.

Notice that as longer chains replace shorter ones, some blocks are discarded along with their contents. Transactions included in these blocks that do not appear in the replacing chain disappear from the ledger. Moreover, if a conflicting transaction exists in the newly adopted chain, the original transaction cannot be included in an extension of the new chain. The mechanism used to choose between different versions of the chain can thus be exploited by a resourceful attacker to reverse payments. This form of attack, as we shall shortly see, is difficult to carry out without access to a large share of computing resources.

Double spending attacks. Consider an attacker that has paid some merchant, has had its transaction embedded in the block chain, but wishes to reverse it (after obtaining some goods in return). The attacker may then use the fact that nodes will adopt an alternative version of the block chain if it is longer than their current one. It can try to create a fork of the current chain that splits off just before the block containing his transaction, and extend this fork in the chain until it is longer than the current chain. Once published, this alternative chain (in which the attacker includes a conflicting transaction that redirects the funds) will take over as the accepted version of events and the attacker's original transaction will be discarded along with the block that contained it.

It is obvious then, that transactions are never fully irreversible in the system; a longer chain may always appear. Such an occurrence, however, becomes increasingly unlikely. Notice that the attacker needs to create enough blocks in his version of the chain to overtake the current main chain. Since block creation requires a difficult proof-of-work computation, the attacker must either have a great deal of computational resources, or be extremely lucky. He must produce blocks at a higher rate than the rest of the network combined in order to overtake the current chain.

Bitcoin's developer Nakamoto has shown in his original analysis that as long as an attacker possesses less than 50% of the computational power in the network, he produces blocks at a lower expected rate than the rest of the nodes, and so the probability of a successful attack on a given transaction decreases exponentially as more blocks are added to the chain on top of it.[25] Each block added is thus considered to add a "confirmation" to all the transactions in preceding blocks as it supports their inclusion in the ledger.

The network is therefore more secure the more computational power there is in the hands of honest nodes. For example, Nakamoto's analysis,[25] (later improved by Rosenfeld[29]), shows that after approximately six confirmations, an attacker with 10% of the computational power can succeed with probability lower than 0.00059. The costs of a mining operation capable of mining even 10% of the blocks is extremely high, establishing a barrier against double spending attacks of transactions that are deeply embedded in the chain.

Merchants—especially those collecting payments in the presence of the buyer—cannot afford to keep customers waiting even for the 10 minutes required for the first confirmation of a transaction. Many have opted instead to accept transactions with 0-confirmations once they have been sufficiently distributed through the network, trusting they will later be included in the block chain. Thus far, relatively few attacks on 0-confirmation transactions have taken place, but such practices still pose a risk.[6,16]

Resilience to the double spending attack relies strongly on the assumption that Bitcoin's P2P network is connected, and that honest nodes are able to communicate. Without communication, blocks and transactions cannot be distributed well. While several mechanisms have been put in place to maintain connectivity, Bitcoin's overlay network has been shown to be susceptible to eclipse attacks in which relatively few attacker nodes manage to attract a sizeable portion of the connections and isolate others from the network.[15] Lower-level infrastructure attacks may also cause problems, especially those committed by adversaries that control many routers, IP addresses, and other network resources (as evidenced by recent BGP hijacking attacks that were successfully used against mining pools[8]).

The 50% attack. A miner that holds over 50% of the network's computational power can create blocks faster than all other nodes combined, and thus represents a more serious threat to the network. It is always able to create blocks at a faster pace than the rest of the network combined, which allows it to change the block chain, and double spend any transaction it issued (regardless of its depth in the block chain). In fact, it is capable of committing much more devastating attacks: by simply generating a chain of empty blocks and ignoring all other blocks it can stop all transactions from entering the block chain. Interestingly, a 50% attacker is still somewhat limited: it cannot move funds it does not control, as transaction messages still must be cryptographically signed by the sender.

Rewards and Incentives

Bitcoin includes an important incentive mechanism that encourages mining activity and thus indirectly increases the system's resilience. Miners are awarded with bitcoins in return for their effort: Each transaction offers a small fee claimed by the node that includes it in a block. Transactions compete for limited space in blocks and so market forces should eventually set the fees. Nodes, in turn, are incentivized to join the network, to collect transactions, and include as many of them as possible in blocks. As a side effect, they contribute their computational power toward improving the network's resilience to double spending.

In addition to fees, creators of blocks are awarded newly created bitcoins. Rewards are issued to the block's creator in a special transaction included in each block called the coinbase transaction. Bitcoin's money creation is gradual, and occurs according to a fixed schedule. Instead of launching the system with all coins in the hands of a single individual, Nakamoto decided to spread their distribution over time. Starting with the Genesis Block, each new block issued 50 bitcoins. The number of bitcoins generated in this manner is halved every 210,000 blocks (approximately every four years), and so the total sum of bitcoins that will ever exist is nearly 21 million. This fixed money supply is one of the key economic features of Bitcoin. It leaves no room for monetary intervention and essentially implies the currency is deflationary (bitcoins may be lost and never replaced if, for example, the private keys needed to transfer them are lost).

Mining bitcoins has become increasingly more attractive as their value has gone up—turning bitcoin mining into a fast-growing industry. Miners have transitioned from using PCs to more efficient hardware such as GPUs, and eventually ASICs—custom designed chips that efficiently perform the operations needed for Bitcoin's proof-of-work. While a single CPU provides several mega hashes per second, modern mining rigs are approximately a million times faster, performing several tera hashes per second. As a result, the network's hash rate has grown to over 300 peta hashes per second, making it more secure against double spending attacks.

Strategic behavior. While rewards have generally attracted more nodes and have strengthened the network, it is important to consider other behaviors nodes may adopt in order to increase their profits from mining. In particular, if nodes find it profitable to deviate from the protocol, the system's performance may deteriorate.

Indeed, some activities related to the basic maintenance of the network have not been properly incentivized

Figure 1. A fork in the block chain.

Two conflicting blocks extend a chain. The two, which are created and held by different nodes, contain somewhat different transaction sets including conflicting transactions (tx9a and tx9b in this example).

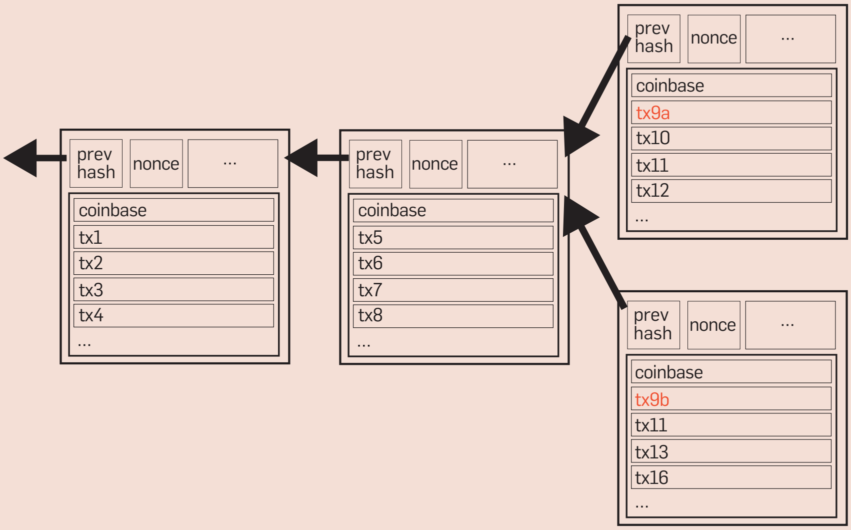
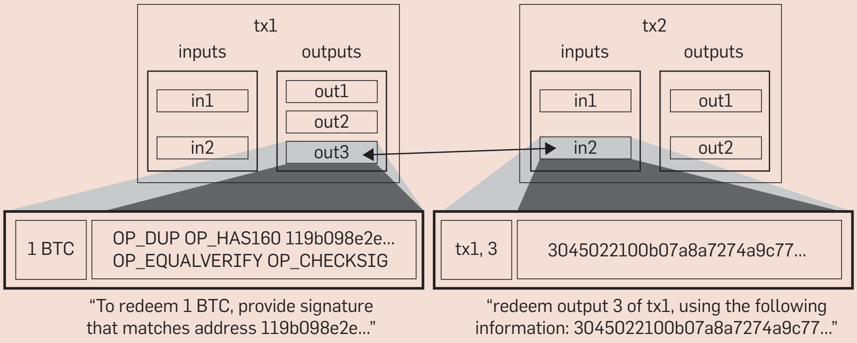


Figure 2. The structure of transactions.


tx1
inputs
outputs
in1
in2
out1
out2
out3
tx2
inputs
outputs
in1
in2
out1
out2
1 BTC
OP_DUP OP_HAS160 119b098e2e...
OP_EQUALVERIFY OP_CHECKSIG
tx1, 3
3045022100b07a8a7274a9c77...
"To redeem 1 BTC, provide signature that matches address 119b098e2e..."
"redeem output 3 of tx1, using the following information: 3045022100b07a8a7274a9c77..."

with payments. Storage costs, for example, are not accounted for and are not reflected in fees. Anyone moving a small amount of money via the block chain creates records that might never be expunged and will forever take up space on all full nodes.

Research has additionally shown that nodes are not properly incentivized to share information. For example, while the protocol requires nodes to flood transaction messages to each other, those who do not distribute messages may gain higher rewards simply by removing the chance that some other node will claim the transaction fees associated with the withheld transaction.[4] Payments to nodes who forward messages may add incentives for distribution and correct this problem.

Another example of bad incentives in the protocol relates to mining behavior: The protocol requires nodes to create blocks that reference the longest chain and to publish them immediately. A paper by Eyal and Sirer[13] shows that nodes can strategically delay the release of blocks they create in order to become more profitable. This selfish behavior is profitable mostly for large miners and is easily detectable, making it somewhat less likely to be used in practice. Still the fact remains: the protocol—in its current form—is susceptible to some level of manipulation by selfish participants.

Mining pools. Given the large number of participants in the network, a small miner can only expect to create a relatively small fraction of the blocks—roughly equivalent to its share of the computational power. For some, this may mean finding a block only after months of continuous attempts. Such rare large rewards may provide a sufficiently high expected payment, but miners that would like to obtain a steady revenue stream also worry about the risk: Payments have high variance and some months may go by without any reward.

The solution to this problem appeared in the form of mining pools: groups of miners that join together and split the profits from mining to receive lower, but more frequent payments. A miner that chooses to participate in a pool contributes his computational power and works to generate the proof-of-work for a block the pool's server is creating. If successful, the pool distributes the block reward among the participants, aiming to provide rewards in proportion to the effort of each miner and guaranteeing them nearly the same expected payment as mining on their own (the difference is due to fees the pool collects for its services and to small inefficiencies in the slightly more complex block creation process). While early implementations of pool reward mechanisms suffered from incentive problems that allowed miners to gain more than their fair share of the rewards through manipulations, newer mechanisms, now in place at most pools, do better.[28] Game theoretic analysis of the competition between pools to attract miners shows there may well be no stable partitions, and that miners will continually switch.[18]

A key aspect of the Bitcoin protocol is the way it represents and changes the ownership of money. Every Bitcoin transaction is in fact a reassignment of money from inputs to outputs.

Pools have been incredibly successful—a relatively small number of them are currently creating the majority of the blocks in the network prompting concerns that too few entities control block creation. CEX.IO, a company that owns a great deal of mining hardware and runs GHash.IO—a popular mining pool—has approached 50% of the network's hash rate on several occasions, a size that could potentially allow them to disrupt the Bitcoin network with a 50% attack. While no attack was launched by the owners of the pool, many have been concerned the system is at risk if, for example, a hacker manages to compromise the pool's servers. The resulting public outcry caused many miners to switch to other pools, and prompted CEX.IO to state they will restrict the size of their pool in the future. It is important to note however, that there is no way to ensure any single entity controls under 50%, as it may be mining using different addresses and through multiple servers. The lack of strong identities in the protocol implies Bitcoin is inherently not secure when a pool controls more than 50% of the computational resources.

The Structure of Transactions

Another key aspect of the Bitcoin protocol is the way it represents and changes the ownership of money. Every Bitcoin transaction is in fact a reassignment of money from inputs to outputs (many to many). Outputs are composed of a sum of money associated with a short script in Bitcoin's scripting language. Scripts can be thought of as a sort of challenge: anyone wishing to move the money associated with an output must provide a string that will make the script return "true."[e] Transaction inputs point to a previous transaction output and contain the string that correctly answers the challenge. Thus, a transaction, which is a combination of inputs and outputs, proves the owner is allowed to transfer money from previous outputs that it owns and redirect it into new transaction outputs. Once an output is used, all of the money associated with it is considered spent. Other transactions that attempt to access the same output will be considered conflicting transactions and will be rejected. Since money from inputs does not necessarily sum up to the amount one may wish to send, transactions often include an output that returns any leftover funds back to the sender.

While it is possible to write simple scripts like "check to see if the input equals 3" that can be easily satisfied, the most commonly used output script is one that requires a cryptographic signature in order to free funds. The script compares the signature provided to the public key associated with a certain bitcoin address and allows the output to be used only if they match (The address is the hash of the public key with some additional bits used for error detection, in case it is mistyped). Ownership of bitcoins is therefore just a matter of knowing the right input to the script. From a practical perspective, a merchant that wishes to receive funds needs to send his address to the buyer. The buyer then creates a transaction with an output that is redeemable by anyone who possesses the corresponding private key, which is kept secret by the merchant for future use.

Other more complicated scripts are also possible; For example, scripts requiring signatures generated by two different keys (effectively implementing a joint account that needs consent from two sources for every transfer).

Privacy, anonymity, and auditability. The structure of transactions and the fact they are publicly available on the block chain allows anyone to follow money and see where it is being moved. This is both a blessing and a curse. On one hand, organizations that wish to do so can reveal which addresses they control and allow anyone to see how they are using their money. On the other hand, the privacy of individuals is compromised. Since addresses are easily and freely generated, it is possible to generate a unique address for every transaction. This helps restore privacy to some extent, but some information is always leaked even when Bitcoin addresses are not reused.[3,20,27]

The mixture of partial privacy and transparency within Bitcoin has led to interesting innovations. The collapse of MtGox, the large bitcoin exchange, which had lost a sizeable amount of bitcoins was followed up by forensic analysis of transaction data that dispelled some possible explanations for its loss of funds.[12] Exchanges have since been pressured to implement mechanisms that allow account owners to securely and privately verify that their balances are indeed held by the exchange. Similar mechanisms have been applied in other domains like crowd funding, online gambling, and charity fund-raising. On the opposite side of the privacy spectrum, some organizations utilize the relative privacy offered by Bitcoin to hide their activities. As an extreme example, criminal organizations like the Silk Road, an online market for illicit goods that had been busted by authorities in the U.S., benefit from the relative anonymity of Bitcoin addresses.

The public aspects of money also enable the use of taint analysis: coins considered to have been involved in illegal activity can be tracked no matter how many times they change hands, and can be treated differently: exchanges, for example, may refuse to accept them. Marking money in this way may have devastating consequences on its fungibility—another important property of money.[24] Issues of privacy, anonymity, taint, and regulation are at the center of debate within the Bitcoin community, and are naturally of great concern to policymakers.

Mixing services and protocols such as CoinJoin allow users to mask the origins and destination of payments by mixing together many transactions and splitting the outputs in ways that do not allow them to be easily associated with the corresponding inputs. Zerocash, a protocol modification designed to provide enhanced anonymity, uses advanced cryptographic tools to allow nodes to process transactions without knowing the details of the transfer.[7] These modifications and others reshape the mixture of privacy and transparency that Bitcoin and similar protocols may provide.

What Does the Future Hold?

Scalability. The Bitcoin protocol is highly wasteful. A high amount of effort is expanded in arbitrary proof-of-work computations. Thus far, no provably secure replacement that uses fewer resources or utilizes the computation for useful purposes has emerged, although many have tried to suggest alternative designs. In addition to the proof-of-work, Bitcoin's design requires wasteful replication. All relevant information is saved at all mining nodes, messages are essentially broadcast through the network, and verification is always repeated. For these reasons, it appears the system would not scale well. Bitcoin's block size has been artificially (and somewhat arbitrarily) limited to 1MB per block. The protocol currently processes under two transactions per second on average, a rate that has been steadily, albeit slowly, increasing. Fortunately, the average transaction size is relatively small, averaging approximately 0.5KB per transaction, which currently allows all transactions generated between block creation events to clear. Concern about the growth of transaction rates has caused some core developers to push for an increase in the block size limit and has sparked lively debate. Those who oppose argue the costs of running nodes will increase beyond the reach of "regular users."

But can Bitcoin scale to process much more significant volumes? As a hypothetical scenario, one may consider rates of 2,000 transactions per second (which are closer to the order of

[e] Complexities related to infinite loops and to the halting problem have been avoided by making the scripting language less expressive. It is not Turing complete.

magnitude of Visa's worldwide transaction volume). With 0.5KB per transaction, the flow of data needed to keep up with all transactions is only approximately 1MB per second (additional protocol messages will in fact require a bit more). Storing all these transactions in the block chain implies storage will grow at a rate of around 2.5TB per month.[f] While this is a high rate of growth, outside the reach of home users, it is certainly manageable for a small mining business even with today's technology.

It is important to note that even mining nodes do not have to hold the entire history of transactions. Some of the contents of the block chain can be safely erased. Furthermore, everyday users of the currency that do not engage in mining do not need to store the full block chain. They can manage with a much smaller portion of the data. Simplified Protocol Verification clients (SPV), also known as light nodes, allow users to connect to the network and download only the information they require. Such nodes are light enough to run on mobile devices, and drastically alleviate storage costs for small users. Miners and others who run full nodes are the only ones that need to hold a full copy of the block chain.

The size of blocks, however, has other important implications. Large blocks take longer to transmit and to propagate through the network. As a result, more conflicting blocks will be created. This fact has been empirically observed,[11] and has severe implications to Bitcoin's resilience to double spending attacks. With many conflicting blocks, the block chain does not grow efficiently, and many of its created blocks are discarded implying that weaker attackers can overtake it. Security deteriorates well before bandwidth limits are reached. An alternative to the longest-chain selection rule, nicknamed GHOST, has been shown to alleviate this security problem.[30] Additional ideas, such as replacing the block chain with a directed acyclic graph structure[19] to include transactions from off-chain blocks, block compression, and off-chain transactions channels[9] offer further improvements to transaction throughput.

Another problem encountered under high transaction rates is the reward distribution between miners becomes skewed in favor of larger, better connected miners (that is, miners connected to the Bitcoin network). This may endanger Bitcoin's decentralized nature, as small miners that cannot invest heavily in connectivity quickly become unprofitable.

Decentralization at risk. While the Bitcoin protocol is decentralized, the current system is in fact controlled in many aspects by small groups of miners, and wallet providers. Protocol development too, is in the hands of relatively few developers.[14]

The race for advanced ASICs used in bitcoin mining is still ongoing, and hardware is often made obsolete within months. As the electricity costs of running a PC far exceed the rewards it generates, mining using CPUs has quickly become a losing proposition. Mining is thus gradually shifting to the hands of larger organizations that continuously invest in the latest hardware. Some have suggested using alternative proof-of-work procedures that will make specialized hardware less effective and will thus weaken this effect (one such example appears in Miller et al.[21]).

Other strong economic forces are also pulling Bitcoin in the same direction of increased centralization. Large miners enjoy effects of increasing returns to scale; They can produce their own hardware, or purchase it en masse, or they may better optimize the location of their mining centers in order to gain access to cheaper electricity. Similar advantages result from the specific nature of the protocol. Storage and bandwidth costs, for example, are the same regardless of the miner's size. This greatly benefits large miners that pay less per generated block for these overheads. Smaller, less profitable competitors, are then slowly eliminated from the market.

Development and protocol changes. Protocol updates in Bitcoin are difficult. Unlike more conventional software, a bug in Bitcoin's core may cause inconsistencies between different versions of the code and may cause the block chain to split. Such an event occurred in March of 2013. A bug in the code caused two versions of the protocol to behave differently (one version refused to accept a block created by the other) which resulted in a long-lasting fork in the block chain. Large mining pools were quickly asked to downgrade to the older version, which eventually resolved the split.

Similarly, intentional updates that do not maintain backward compatibility may cause the chain to fork if they are not accepted by all. Such updates cannot be rolled out gradually—they require a majority of the network to accept them before they are activated. Bitcoin's core developers have therefore been extremely conservative with updates. This justifiably careful behavior also implies the protocol itself is slowly "calcifying," as substantial updates become progressively more difficult to roll out.

Alternative currencies. Bitcoin's open source code has been used to launch many alternative currencies (altcoins). Many have been created by applying relatively minor modifications to its code. One example is Litecoin, which aims to be "the silver to Bitcoin's gold." Litecoin's proof-of-work hashing algorithm has been changed in hope of preventing ASICs from dominating the mining race and its blocks are created at a somewhat accelerated rate of once every 2.5 minutes (ASICs were eventually developed for Litecoin mining as well). Many alternative currencies have found some following (for example, Dogecoin is based on a famous Internet meme), but have usually struggled to attract many miners, and to maintain a secure network.

Not all altcoins are minor modifications. Some include more substantial changes, and have taken ideas from Bitcoin to new realms. Namecoin, for example, uses its block chain as a key-value store rather than to manage a currency (one of its uses is as a distributed alternative to DNS).

In this context, it is also worthwhile to mention Ripple and Stellar,[26] (https://www.stellar.org) two companies developing protocols not derived directly from Bitcoin, but that create a distributed system for money transfer. Both are primarily based on a network of IOUs that are transferred locally, and have dif-

[f] The block chain's size today is approximately 40GB, and it currently includes all of the transactions since Bitcoin was launched in 2009.

ferent consensus mechanisms.[26]

Altcoins are considered by many to be the proving grounds for new risky ideas that may someday be incorporated into Bitcoin if they have proven to be viable. Others complain that there are many "pump-and-dump" schemes, that is, coins that are created with a lot of hype to lure in naïve speculators who invest money and get little in return. Ideas like sidechains[5] that may allow bitcoins to be exchanged for other compatible alternative coins have been suggested as potential mechanisms that allow for easier integration with Bitcoin, and may be the path for safer innovation in cryptocurrencies.

Beyond money. While initially designed only to encode monetary transfers, it quickly became clear that Bitcoin's block chain can be used to encode other pieces of information.

Examples range from innocuous ASCII art images to WikiLeaks cables that have been embedded in transactions. This has raised several concerns both regarding legal aspects of embedding copyrighted or otherwise prohibited information into the block chain (which is then copied to every full Bitcoin node).

Discontent with the scripting capabilities that Bitcoin offers, some higher-level protocols have opted to extend the functionality of its scripting language to include additional actions. Counterparty, Omni, and Colored-Coins are several higher-level protocols that do just that. Reasoning that Bitcoin's network is large and highly secure, these protocols use Bitcoin transactions to encode more sophisticated scripts that allow for multiple currencies to exist within its block chain. Other higher-level functions like distributed exchanges, bets, and financial derivatives are also enabled.

The Ethereum project, which uses a separate block chain, has taken transaction scripts one step further and developed a Turing-complete scripting language. Ethereum allows anyone to create contracts, which are essentially programs that are executed jointly by the nodes in the Ethereum network. Ethereum's block chain is used to maintain the state of each contract, and transaction messages generate events that update these states.

The realization that decentralization has value in and of itself, as well as the rise of platforms like Ethereum, has led some to believe computer programs can become fully autonomous economic entities. "Decentralized Autonomous Corporations" (DACs), can collect fees (paid with cryptocurrencies) for services rendered, and use them to pay for servers, and other resources they consume. Existing within decentralized platforms, they can truly have a life of their own, independent of their creator, without depending on any single machine to run their code.

Conclusion

Bitcoin's design fundamentally reshapes and reimagines money—one of humanity's most basic and foundational social constructs. Essentially allowing us to transmit value over the Internet just as easily as we transmit information, its disruptive nature promises to change markets, enable new business models, and impact the ability of governments to control money and to regulate businesses. While still facing many challenges, a steady stream of innovations and solutions is continuously being developed to address its shortcomings. The evolutionary path of the protocol and of the system itself is greatly influenced by the protocol's technical strengths and weaknesses, but also by strong social, political, and economic undercurrents. Miners, developers, regulators, and adopters all affect the direction of its growth. With ongoing development, and possible applications beyond the financial domain, Bitcoin, and other protocols that extend it, may yet come to deeply impact our lives.

References

1. Aaronson, S. Quantum copy-protection and quantum money. In *Proceedings of the 24th Annual IEEE Conference on Computational Complexity*. IEEE, 2009, 229–242.
2. Anderson, R. and Murdoch, S.J. EMV: Why payment systems fail. *Comm. ACM, 57*, 6 (June 2014), 24–28.
3. Androulaki, E., Karame, G.O., Roeschlin, M., Scherer, T. and Capkun, S. Evaluating user privacy in Bitcoin. *Financial Cryptography and Data Security*. Springer, 2013, 34–51.
4. Babaioff, M., Dobzinski, S., Oren, S. and Zohar, A. On bitcoin and red balloons. In *Proceedings of the ACM Conf. on Electronic Commerce*. ACM, 2012, 56–73.
5. Back, A. et al. Enabling blockchain innovations with pegged sidechains; http://blockstream.com/sidechains.pdf.
6. Bamert, T., Decker, C., Elsen, L., Wattenhofer, R. and Welten, S. Have a snack, pay with bitcoins. In *Proceedings of the 13th IEEE International Conference on Peer-to-Peer Computing*, Sept. 2013.
7. Ben-Sasson, E. et al. Zerocash: Decentralized anonymous payments from Bitcoin. In *Proceedings of the IEEE Security and Privacy Symposium*. IEEE, 2014.
8. BGP hijacking for cryptocurrency profit; http://www.secureworks.com/cyber-threatintelligence/threats/bgp-hijacking-for-cryptocurrencyprofit/.
9. Bitcoin lightning network; https://lightning.network/.
10. BitLicense; http://www.dfs.ny.gov/legal/regulations/adoptions/dfsp200t.pdf.
11. Decker, C. and Wattenhofer, R. Information propagation in the Bitcoin Network. In *Proceedings of the 13th IEEE International Conference on Peer-to-Peer Computing* (Sept. 2013).
12. Decker, C. and Wattenhofer, R. Bitcoin transaction malleability and MtGox. In *Proceedings of the 19th European Symposium on Research in Computer Security* (Wroclaw, Poland, Sept. 2014).
13. Eyal, I. and Sirer, E.G. Majority is not enough: Bitcoin mining is vulnerable. *Financial Cryptography and Data Security, LNCS* (2014). 436–454.
14. Gervais, A., Karame, G., Capkun, S. and Capkun, V. Is Bitcoin a decentralized currency? *IACR Cryptology ePrint Archive 829* (2013).
15. Heilman, E., Kendler, A., Zohar, A. and Goldberg, S. Eclipse attacks on Bitcoin's peer-to-peer network. In *Proceedings of 24th USENIX Security Symposium*. Aug. 2015 (to appear); http://eprint.iacr.org/2015/263.pdf.
16. Karame, G., Androulaki, E. and Capkun, S. Two bitcoins at the price of one? Double-spending attacks on fast payments in Bitcoin. *IACR Cryptology ePrint Archive 248* (2012).
17. Lamport, L., Shostak, R. and Pease, M. The byzantine generals problem. *ACM Trans. Programming Lang. Systems 4*, 3 (1982), 382–401.
18. Lewenberg, Y. Bachrach, Y. Sompolinsky, A. Zohar, and Rosenschein, J.S. Bitcoin mining pools: A cooperative game theoretic analysis. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. 2015, 919–927.
19. Lewenberg, Y., Sompolinsky, Y. and Zohar, A. Inclusive block chain protocols. *Financial Cryptography and Data Security*. Springer, 2015 (to appear).
20. Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G.M. and Savage, S. A fistful of bitcoins: Characterizing payments among men with no names. In *Proceedings of the 2013 Conference on Internet Measurement*. ACM, 127–140.
21. Miller, A., Juels, A., Shi, E., Parno, B. and Katz, J. Permacoin: Repurposing Bitcoin work for data preservation; http://cs.umd.edu/ amiller/permacoin.pdf, 2014.
22. Miller, A. and JLaViola, Jr., J.J. Anonymous byzantine consensus from moderately hard puzzles: A model for Bitcoin, 2014.
23. Moore, T. and Christin, N. Beware the middleman: Empirical analysis of Bitcoin-exchange risk. *Financial Cryptography*. Springer, 2013, 25–33.
24. Möser, M., Böhme, R. and Breuker, D. Towards risk scoring of Bitcoin transactions. *Financial Cryptography and Data Security, Lecture Notes in Computer Science*. Springer, 2014, 16–32.
25. Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. 2008; https://bitcoin.org/bitcoin.pdf.
26. Ripple; https://ripple.com/.
27. Ron, D. and Shamir, A. Quantitative analysis of the full Bitcoin transaction graph. *Financial Cryptography and Data Security*. Springer, 2013, 6–24.
28. Rosenfeld, M. Analysis of Bitcoin pooled mining reward systems. arXiv preprint (2011); arXiv:1112.4980.
29. Rosenfeld, M. Analysis of hashrate-based double spending; https://bitcoil.co.il.Doublespend.pdf, 2012,.
30. Sompolinsky, Y. and Zohar, A. Secure high-rate transaction processing in Bitcoin. *Financial Cryptography and Data Security*. Springer, 2015 (to appear).
31. Wiesner, S. Conjugate coding. *ACM SIGACT News 15*, 1 (1983) 78–88.

Aviv Zohar (avivz@cs.huji.ac.il) is a senior lecturer in the Rachel and Selim Benin School of Computer Science and Engineering, at The Hebrew University of Jerusalem, Israel, and a visiting researcher at Microsoft Research, Israel.