---

slug: formalization-mev title: On the Formalization of MEV authors: [alejo] tags: [research] image: https://i.imgur.com/mErPwqL.png hide_table_of_contents: false forum_link: https://collective.flashbots.net/t/on-the-formalization-of-mev/879

Thanks to Phil Daian, Alex Obadia, and Mahimna Kelkar for plenty of discussions on the topic.

Since its introduction in the Flashboys 2.0 paper of 2019 by Daian *et al.*, a lot has been said about *Miner* (now *Maximal*) *Extractable Value*, or MEV. In particular, the launch of the Flashbots Auction propelled what is today a billion dollar economy across various blockchains and centralized exchanges. From thrilling Twitter threads to academic research papers, the MEV phenomenon has captured a central spot in the cryptocurrency discourse. Oddly, however, there is no agreed upon formal definition of MEV.

While some may argue that the widely shared, intuitive notion of MEV will be sufficient in most scenarios, we believe that proper formalization is critical to establishing a foundation upon which complex theorizing can take place. As Tim Roughgarden put it in a recent talk on building a theory of DeFi, the first step before "easy" and later "difficult" theorems is to have definitions and basic vocabulary. Further, as was evidenced by a recent public discussion where some claimed that arbitrage is not MEV, it might even be the case that we don't share an intuitive notion of MEV after all! A unifying formal definition of MEV would certainly help.

As it turns out, however, formalizing MEV in a robust, general way is no easy task. In this post, we explore some of the difficulties we encounter when trying to come up with such a definition. We start by reviewing some of the existing formalizations, point out some of their problems, and go on a quest trying to amend some of them. While we present new definitions that improve on some of these issues, our main contribution rests in highlighting many of the subtleties involved, paving the way for a more systematic approach in future work around MEV.

Current MEV definitions

The original Flashboys paper defines MEV as *"the total amount of Ether miners can extract from manipulation of transactions within a given timeframe, which may include multiple blocks' worth of transactions"*, but stays short of attempting a formal definition. Most recently, the widely adopted working definition rings something like:

> *MEV is the value that can be permissionlessly extracted by block proposers by reordering, censoring, or inserting transactions.*

Perhaps the definition that comes closest to formalizing this is the one given in the recent Clockwork Finance paper via the following two expressions:

$$ \mathsf{EV}(p,B,s)=\max_{(B_1, \ldots, B_k)\in B} \left\{ b(p, s_k)-b(p,s)\right\} \tag{1} $$

and:

$$ k\mathsf{-MEV}(p,s)=\mathsf{EV}(p, \mathsf{validBlocks}_k(p,s), s). \tag{2} $$

Here, EV is the *extractable value* by player p in state s given a set of valid block sequences B, $(B_1, \ldots, B_n)$ is one such sequence, and $b(p, s_k)$ is the balance of player p in the state resulting of applying blocks $(B_1, \ldots, B_k)$ to s. *k*-MEV is the *k-maximal extractable value* by a player p in state s acting as block proposer, where $\mathsf{validBlocks}_k$ is the set of all valid block sequences of k blocks that p can create, and single-block MEV is just 1-MEV.

These expressions were slightly tweaked from those in the paper for notational simplicity, but are otherwise equivalent. In particular, we consider balances of players as opposed to accounts (omitting a sum over accounts controlled by a player), and remove the explicit reference to the chain's native asset; we will return to this point later.

We will use this definition of MEV as a starting point, noting that most other papers provide similar definitions that face the same limitations, or do not provide formal definitions at all.

Existing limitations

We start by noting a fatal flaw in the above expressions: the maximal extractable value depends on the player p! This means that if p has, say, some pending airdrop claim, their MEV will be larger than for a player who doesn't. While this might make sense for the extractable value, it is certainly at odds with the idea of a "permissionlessly extracted" value.

Upon closer inspection, it is not entirely clear what the notion of "player" is actually referring to. We can identify at least three intertwined meanings: i) a player as a transaction signer, having balances and controlling accounts, ii) a player as an actor in the protocol game, having (or lacking) block proposing rights, and iii) a player in the networking sense, a node operator affected by latencies and having a unique mempool view.

While the latter meaning does probably not apply in this formulation (although we will come back to it later), meanings i) and ii) are somewhat conflated: p certainly refers to i) when talking about balances in (1), but in going from (1) to (2) we also ascribe p block proposing rights, in line with meaning ii). We argue that a proper definition of MEV *should be independent of the player in the sense of i)*, that is, it should not depend on particular signing rights. With respect to ii), we will define MEV *given* block proposing privileges. This effectively decouples the problem into value extraction on one hand, and obtaining sequencing rights on the other, which might prove useful when thinking about extraction costs, network security, etc.

Other caveats with the above definitions are the treatment of multi-block MEV (related to the tangling of meanings i) and ii) above), the omission of fees arising from reverted transactions, and the inadequacy of the notion of blocks when attempting to generalize MEV to the cross-domain setting. In what follows, we attempt to patch the definitions to solve some of these issues when possible, and discuss some other difficulties we find along the way.

Patching MEV

As mentioned above, the first task is to come up with a truly permissionless MEV definition. We will keep the player dependency in the definition of extractable value, but get rid of it when moving to MEV. We note here that we use *player* in sense i) above, giving it, for both EV and MEV, full block sequencing rights. We propose the following:

$$ \mathsf{EV}(p, s)=\max_{B \in \mathsf{validBlocks}(p)} \left{ b(p, B(s))-b(p,s)\right} \tag{3} $$

$$ \mathsf{MEV}(s)=\min_{p\in P}{\mathsf{EV}(p, s)}. \tag{4} $$

The first expression here closely resembles (1), but we removed the dependency on the set of valid block sequences, which is implicit, and we're only considering a single block (more on this later). Here $\mathsf{validBlocks}(p)$ is the set of valid blocks that can be proposed by p ($\mathsf{validBlocks}_1(p,s)$ before, omitting the amount of blocks and state dependency for succinctness). $B(s)$, in turn, denotes the state obtained by applying block B on top of state s.

In expression (4), we obtained a definition for MEV that, as desired, is independent of the player (denoting P the set of players). While it might perhaps be counterintuitive to find a minimum in the definition of *maximal* extractable value, this minimum simply encodes the idea that extraction should be permissionless. EV already takes care of maximizing, the value extractable permissionlessly is the one the *least privileged actor* can take from the network (again, assuming they have block proposing rights).

This definition begs the question, however, what happens when extraction requires upfront capital? Definition (2) did not have this problem since it had an explicit dependency on the player, but having now removed it, we need to take into account that some MEV might only be extractable at certain levels of initial capital. We note however that gas fees are not part of the requirement here, since a proposer can sequence "free" transactions at will, so that in general MEV might be greater than zero even with no initial capital.

Still, we want to make the dependency on capital explicit, since many MEV opportunities depend on it. We write (using EV from (3)):

$$ \mathsf{MEV}(s;K)=\min_{{p\in P | b(p,s)\geq K}}{\mathsf{EV}(p, s)}. \tag{5} $$

This definition tells us that the maximal extractable value in state s for initial capital K is the value that can be extracted by any player that possesses at least that amount of initial capital.

The next step we consider is what happens with transactions in the mempool. In the above we were considering "valid blocks", but crucially these can contain reverted transactions, that pay fee, but do not modify the state. This is a tricky point, as it involves meaning iii) above for the player due to the fact that different views of the mempool yield different sets of valid blocks. While in practice searchers extracting MEV are constantly looking at the mempool for opportunities, transactions will eventually need to be included in a block to modify the state and give rise to the opportunity, so there is no loss of generality in terms of valid transactions if we only consider state changes as opposed to the more general notion of valid blocks. We do lose reverted transactions as a source of MEV in this case, so we could try to modify our formulas to include a view of the mempool dependent on the player, but this would conflate meanings i) and iii), and we would run into trouble when minimizing over players. Considering that the mempool architecture is particular only to some domains, this would also limit the generalizability of the expressions. We thus explicitly leave out reverted transactions as a source of MEV, but note that they are part of the revenue that sequencers take home, and contribute to the negative externalities of MEV extraction, as quantified by the *extractable value cost*.

Leaving out reverted transactions enables us to go even further, beyond the notion of blocks, which will enable us to consider MEV in more general domains, like centralized exchanges. We rewrite our definition for extractable value as:

$$ \mathsf{EV}(p, s)=\max_{\left{s'\in S|s\xrightarrow{a_p} s'\right}} \left{ b(p, s')-b(p,s)\right}. \tag{6} $$

Here, S is the set of all states, and the notation $s\xrightarrow{a_p} s'$ means that state s' is *reachable* from state s by some action or sequence of actions a_p by player p. Together with equation (5) above, we achieved defintions that will allow us to easily generalize to the cross-domain case (see below), and solves most of the issues we encountered with definitions (1) and (2).

Outstanding issues and extensions

In patching the MEV definition, we moved to single blocks, sweeping under the rug the issue of multi-block MEV. In fact, this is automatically taken into account by our latest expressions (5) and (6), since by expressing EV in terms of state as opposed to blocks, the formulas apply to whatever period the proposer has ordering rights in. The question now becomes one of how to get those ordering rights. In order to do this cleanly we need to ascribe probabilities for the different events (say producing a single block, two consecutive blocks, etc.), so we can come up with an expected value for the total MEV. This, however, is beyond the scope of the MEV formalization, since expression (5) can be simply plugged in once the adequate ensemble is defined.

Another topic we touched upon in passing is that of *cross-domain* MEV. In a world where different chains (or more generally *domains*) have their own mechanisms for state updates, but are effectively linked by dependencies in their states (think an L1 deposit affecting L2 balances when processed), we expect to find MEV that can be only extracted by sequencing state changes jointly in more than one domain. Our formulation in terms of states is amenable to this extension, with the caveat that different domains have different native assets, which we need to take into account. We won't go into the details here, but this can be tackled introducing *pricing functions* for translating from one domain to the next. At first approximation, we can take a pricing function $p_{i\rightarrow j}$ to go from the native asset in domain i to the one of domain j, and ask that $p_{j\rightarrow i}=1/p_{i \rightarrow j}$. More realistically, we expect price to be a player-dependent function of many factors as the different volumes of the assets in the different domains, the trust assumptions of the domains, etc.

We note that, throughout, we considered EV and MEV as *revenue* to the player, never taking into account the costs. This bodes well with our definitions in terms of players with given ordering rights, since getting those rights is arguably the costliest component of MEV extraction (although compute cost might be non-trivial considering the ordering problem is an NP-complete knapsack problem). In any case, it seems cleaner to think of MEV as only a revenue component, and consider the extraction costs separately. Like in the multi-block setting, we can define a probabilistic ensemble for obtaining the ordering rights, and consider the costs associated to each probability distribution. It is trickier, however, in that ordering rights are typically granted in units that are characteristic of each domain (e.g. propose one block), while costs are typically expressed as rates (per unit of time). So while MEV will usually come in, say, blocks, the cost of producing those blocks will be expressed in units of time, and the specifics of their relation will be particular to each domain. How each domain achieves finality will also be critical to establishing this relation (perhaps there was a great MEV opportunity 1 year ago, but the cost of re-orging the chain to get it would be prohibitive), so we don't expect a general formulation of MEV to accommodate it.

Finally, we have only considered "sure-thing" MEV by expressing it in terms of balances that increase after a state change (for the MEV to be positive). This falls short of describing the more general notion of "probabilistic MEV", where actors are comfortable taking risk in expectation of later rewards. Examples of this are buyouts of new token listings in expectation of rising prices, or more esoterically front-running of NFT bids. It is likely that most of these opportunities can be described by incorporating pricing functions, and we look forward to seeing work in this direction.

Conclusion

The burgeoning phenomenon of MEV begs for a consistent formal approach to unlock proper theory to emerge (think automated auditing of MEV exposure of smart contract systems for an example, in line with the work presented in the Clockwork Finance paper cited above). Formalizing MEV, however, involves droves of technicalities that usually trade off generality for completeness (like the case of whether to include reverted transactions). Here, we provided a clear meaning to the notion of players, which led us to separate the problem of achieving sequencing rights from that of value extraction, allowing us to go to great generality in defining MEV. Our formulas (5) and (6) provide a consistent, easily generalizable definition that can be taken to the multi-domain world. We also highlighted many of the issues faced in achieving formal definitions of MEV, which we hope will allow for a more systematic treatment of the subject.