# Notes on Myerson's Revenue Optimal Mechanisms

Hu Fu

October 29, 2017

These notes contain exposit aspects of Myerson's revenue optimal mechanism. The first part explains a natural thought process that would lead one from the characterization of revenue (as virtual surplus, Lemma 1) to the procedure of ironing, all from first principles. The second part is a sketch of the steps taken to derive Myerson's auction (with ironing) using revenue curves.

## 1 Setting and Notations

We have a single item to sell, and each bidder $i$'s value $v_i$ is drawn independently from a knwon distribution whose cumulative density function is $F_i$ with derivative $f_i$.

We denote by $x_i(v_i, v_{-i})$ the allocation to bidder $i$ when the bid/value profile is $v_i$ and $v_{-i}$, and $p_i(v_i, v_{-i})$ the payment made by bidder $i$. We use $x_i(v_i)$ to denote the *interim* allocation when bidder $i$'s value is $v_i$, i.e., $x_i(v_i) = \mathbf{E}_{v_{-i}}[x_i(v_i, v_{-i})]$, and similarly for $p_i(v_i)$.

## 2 Ironing

In this section we explicate a natural thought process that leads one to the precise form of ironing. Section 3 can be largely seen as another way of doing this. The starting point of the thought process here is Lemma 1, the characterization of revenue as virtual surplus, whereas Section 3 departs from Myerson's original proof at an earlier point.

**Lemma 1** (Myerson, 1981)**.** *A mechanism is Bayesian incentive compatible (BIC) if and only if for each bidder $i$, $x_i(\cdot)$ is a non-decreasing function, and the payment $p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(t)\,\mathrm{d}t$. Furthermore, the expected revenue of a BIC mechanism is $\sum_i \mathbf{E}_{v_i}[x_i(v_i)\varphi_i(v_i)]$, where $\varphi_i(v_i) = v_i - \frac{1-F_i(v_i)}{f_i(v_i)}$ is the virtual value of $v_i$.*

A distribution $F_i$ is said to be *regular* iff $\varphi_i(v_i)$ is a non-decreasing function of $v_i$. When all bidders' value distributions are regular, it is straightforward to see that allocating to the bidder with the highest nonnegative virtual value maximizes the the expression

$$\text{Rev} = \sum_i \mathbf{E}_{v_i}\left[x_i(v_i)\right] = \sum_i \mathbf{E}_{\boldsymbol{v}}\left[x_i(\boldsymbol{v})\right],$$

and the resulting allocation rules are obviously monotone. However, when the distributions are not regular, this may not produce an incentive compatible mechanism. In order to remedy this, Myerson introduced the *"ironing"* procedure. We explain the intuition. The goal is to transform an irregular distribution so that regularity is restored, while any monotone allocation rule is guaranteed to generate no more revenue with respect to the original distribution than with respect to the new distribution.

**Lemma 2.** *For any bidder $i$, let $[a,b]$ be an interval on which $\varphi_i(\cdot)$ is non-increasing, and $x_i(\cdot)$ be the allocation rule of any Bayesian incentive compatible mechanism. Then the allocation rule*

$$x_i'(v_i) = \begin{cases} x_i(v_i), & v_i \in [0,a) \cup (b, +\infty) \\ \frac{\int_a^b x_i(t)\,\mathrm{d}t}{F_i(b)-F_i(a)}, & v_i \in [a,b]. \end{cases}$$

*is a feasible, monotone allocation rule, and its revenue is no less than that of $x_i(\cdot)$.*

*Proof.* The new allocation rule "pools" the types in $[a,b]$, and gives them the average allocation on the interval. The monotonicity of $x_i'(\cdot)$ is immediate from the monotonicity of $x_i(\cdot)$. That it generates more revenue is quite intuitive. We give a formal argument: for $v_i \notin [a,b]$, $x$ and $x_i'$ are identical and hence generate the same virtual surplus; on $[a,b]$, let $F_i[\cdot|v_i \in [a,b]]$ denote the conditional distribution of $v_i$ given $v_i \in [a,b]$ (which simply has density $f_i/(F_i(b)-F_i(a))$),

$$\begin{aligned}\int_a^b x_i'(v_i)\varphi_i(v_i)f_i(v_i)\mathrm{d}v_i &= \frac{\int_a^b x_i(v_i)f_i(v_i)\,\mathrm{d}v_i}{F_i(b)-F_i(a)} \cdot \int_a^b \varphi_i(v_i)f_i(v_i)\,\mathrm{d}v_i \\ &= \mathbf{E}_{v_i\sim F_i[\cdot|v_i\in[a,b]]}\left[x_i(v_i)\right] \cdot \mathbf{E}_{v_i\sim F_i[\cdot|v_i\in[a,b]]}\left[\varphi_i(v_i)\right](F_i(b)-F_i(a)) \\ &\geq \mathbf{E}_{v_i\sim F_i[\cdot|v_i\in[a,b]]}\left[x_i(v_i)\varphi_i(v_i)\right](F_i(b)-F_i(a)) = \int_a^b x_i(v_i)\varphi_i(v_i)f_i(v_i)\,\mathrm{d}v_i.\end{aligned}$$

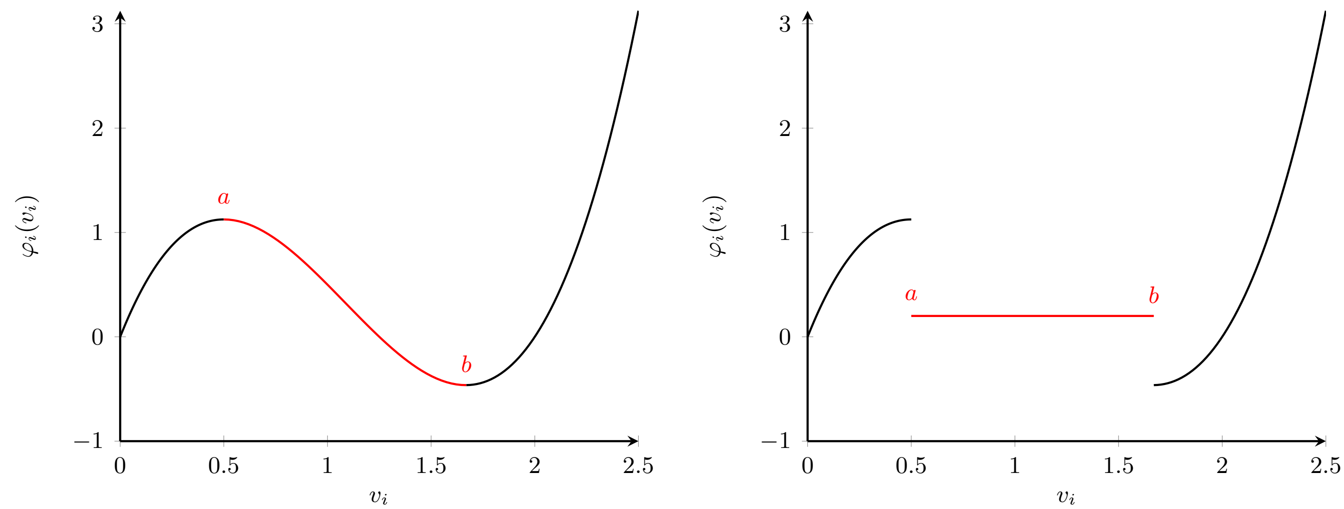
The inequality is an application of Harris inequality, which states that for any non-decreasing function $f$ and non-increasing function $g$ on $\mathbb{R}$ and any probability measure on $\mathbb{R}$, $\mathbf{E}[fg] \leq \mathbf{E}[f]\,\mathbf{E}[g]$.[1] It is here that we make use of the assumption that $x_i$ is monotone non-decreasing and $\varphi_i$ is monotone non-increasing on $[a,b]$.

To see that $x_i'$ is feasible, we describe a mechanism that implements it, which uses as a black box any mechanism $\mathcal{M}$ that implements $x_i$: for any $\boldsymbol{v}_{-i}$, when bidder $i$ reports $v_i \notin [a,b]$, allocate $x_i(v_i, \boldsymbol{v}_{-i})$; for $v_i \in [a,b]$, randomly draw another value $v_i'$ from $F_i[\cdot|v_i \in [a,b]]$, and allocate $x_i(v_i', \boldsymbol{v}_{-i})$. This mechanism is feasible because it can be also seen as running $\mathcal{M}$ while doing the resampling for bidder $i$. From the point of the view of the mechanism, bidder $i$'s value distribution is still $F_i$; in other words, the resampled distribution and the original distribution are not distinguishable. □

For revenue optimality, Lemma 2 allows us to restrict attention to only mechanisms whose allocation rule $x_i$ is flat on $[a,b]$. But any such mechanism is in fact indifferent if we were to replace each of the virtual values on $[a,b]$ by their (weighted) average, $\mathbf{E}_{v_i\sim F_i[\cdot|v_i\in[a,b]]}[\varphi_i(v_i)]$, as Figure 1b shows.

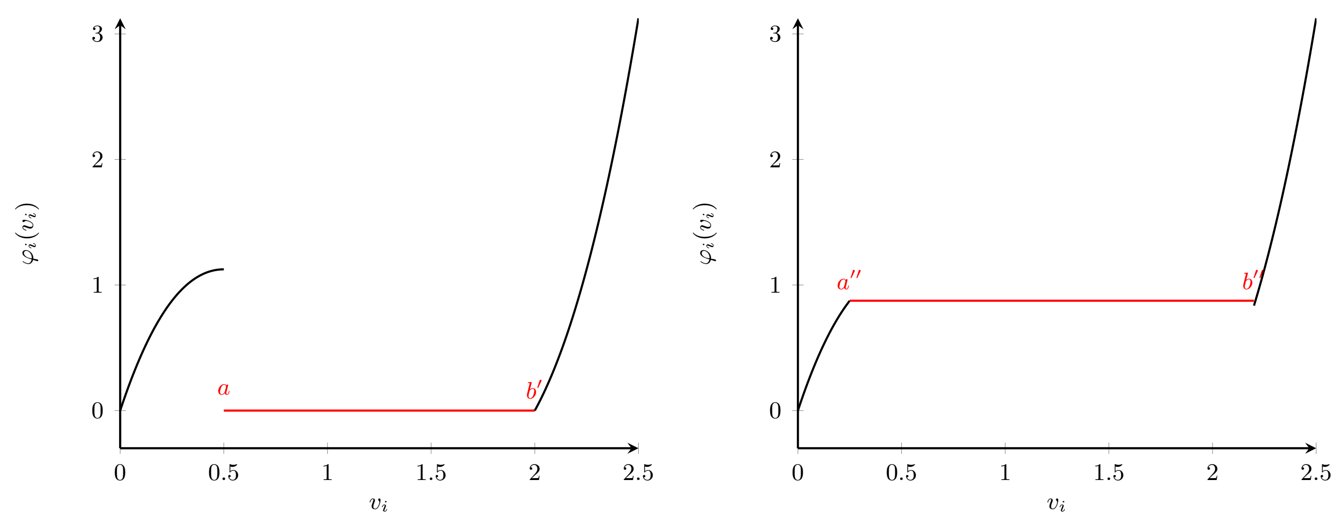
This operation, known as "ironing", flattens the virtual value function on $[a,b]$, but the resulting function is still not monotone, as there is a sharp drop at both points $a$ and $b$. It would be natural to repeat the procedure in Lemma 2: whenever the "next type" after the right end of the ironed region has a virtual value below the flattened area, one may extend the ironed region to include that type; the ironed region's virtual value drops while that of the "next type" rises, so the gap

[1] This fact, which is not hard to prove, is an analog of Chebyshev's sum inequality for discrete domains; the Harris inequality itself is a special case of the FKG inequality.




(a) (b)

Figure 1: A virtual value function with a decreasing interval $[a, b]$ and (naïve) ironing of that interval.




(a) (b)

Figure 2: Extending the ironed region rightward and the final ironed region.

becomes smaller. One may repeat this until there is no drop at the right end of the ironed region. However, this procedure keeps pulling down the ironed region, which only enlarges the drop at $a$, the left end of the ironed region (Figure 2a). So in fact one needs to include the type "immediately preceding" $a$ in the ironed region as well, and one should keep repeating the procedure until there is no drop at the left end of the ironed region. But then the right end now may again see a drop since the flattened area rose.

The question boils down to deciding on an interval $[a'', b'']$, so that when ironing over $[a'', b'']$, the resulting average virtual value is no smaller than what lies on the left of $a''$ and no larger than what lies on the right of $b''$ (assuming the virtual value is monotone elsewhere).

Of course we are under other constraints, as otherwise ironing the whole value space trivially satisfies this. Recall that, in the region-growing procedure, we could assimilate a type immediately preceding or succeding the interval only if there is non-monotonicity at the type (as required by the conditions of Lemma 2). If we denote by $\varphi_i([c, d])$ the average virtual value on $[c, d]$:

$$\varphi_i([c, d]) := \frac{\int_c^d f_i(v_i)\varphi_i(v_i)\,\mathrm{d}v_i}{F_i(d) - F_i(c)},$$

then the constraint on the ironed region $[a'', b'']$ could be stated as:

(i) $\varphi_i(a'') \leq \varphi_i([a'', b'']) \leq \varphi_i(b'')$;

(ii) for any $c, d$ such that $[c, d] \subseteq [a'', b'']$, $\varphi_i(c) \geq \varphi_i([c, d]) \geq \varphi_i(d)$.

How were one to think about these constraints involving every possible pair of values? Let's make an analogy for $\varphi_i([c, d])$ in physics terms. If we see $\varphi_i(v_i)$ as the instantaneous veolocity of a body moving on a straight line and $f_i(v_i)$ as the length of time during which the body moves at this speed, then $\varphi_i([c, d])$ describes the average speed during time from $F_i(c)$ to $F_i(d)$. The constraints on the choice of $a'', b''$ can now be translated in terms of velocities in this new space:

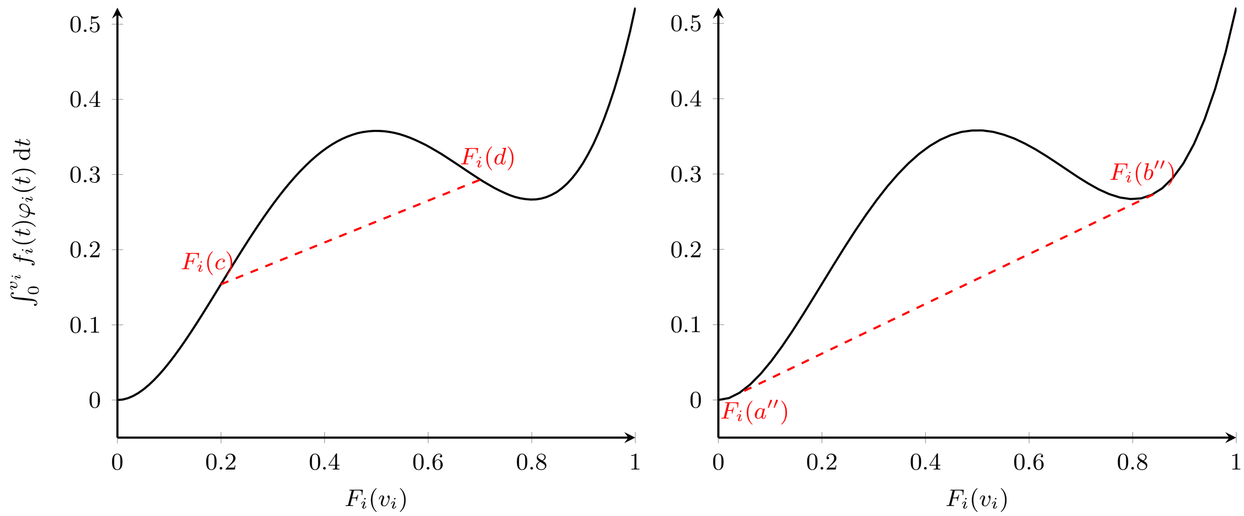
(i') the instantaneous speed at time $F(a'')$ should be no more than the average speed on $[F(a''), F(b'')]$, which in turn should be no more than the instantaneous speed at $F(b'')$;

(ii') for any $[c, d] \subseteq [a'', b'']$, the instantanenous speed at time $F(c)$ should be no less than the average speed on $[F(c), F(d)]$, which is in turn no less than the intantaneous speed at time $F(d)$.

It is now natural to view the problem from the perspective of *distance* traveled, plotted against the *time elapsed*.[2] For each value $v_i$, the total time elasped is $F_i(v_i)$, and the total distance traveled is $\int_0^{v_i} f_i(t)\varphi_i(t)\,\mathrm{d}t$. Figure (3a) shows such a plot. One may call the distance here the *cumulative virtual value*, and the time is naturally the cumulative density.

In this plot, the derivative of the plotted function at $F_i(c)$ is the "instantaneous speed" at this point, which is $\varphi_i(c)$; if we connect by a straight line two points whose $x$-coordinates are $F_i(c)$ and $F_i(d)$ respectively, then the slope of the line is the "average speed over $[c, d]$", which is $\varphi_i([c, d])$. The requirements on the final ironed region $[F_i(a''), F_i(b'')]$ can now be translated as

[2] One may recall the well-known interview question where, given an array of natural numbers, one is asked to construct a linear sized data structure that allows to compute, in constant time, of the sum of elements over any consecutive segment of the array. The simplest answer is to store the partial sum from the first element to the $k$-th, for each $k$. When asked to compute the sum over the segment going from the $i$-th element to the $j$-th, one simply takes the difference between the $j$-th and the $(i-1)$-st stored partial sum.




(a) The slope of the line connecting $F_i(c)$ and $F_i(d)$ is $\varphi_i([c,d])$.

(b) The line connecting $F_i(a'')$ and $F_i(b'')$ constitute part of the convex hull of the plot of the "cumulative virtual value".

Figure 3: Taking the convex hull of the "cumulative virtual value" plotted against the cumulative probabilities.

(i") When replacing the function on $[F_i(a''), F_i(b'')]$ by a straight line connecting the two ends, we should see a convex function (assuming the function was originally convex to the left of $F_i(a)$ and to the right of $F_i(b)$;

(ii") For any $c \in [a'', a]$ and $d \in [b, b'']$, the derivative of the distance function at $F_i(c)$ should be no less than the slope of the line connecting $F_i(c)$ and $F_i(d)$, which in turn should be no less than the derivative at $F_i(d)$.

If we assume $\varphi_i(\cdot)$ to be monotone on $[0, a)$ and $(b, +\infty)$, that is, the "cumulative virtual value" function plotted in Figure (3a) is convex on $[0, F_i(a))$ and on $(F_i(b), 1]$, then it is not hard to see that the above two requirements stipulate that $[F_i(a''), F_i(b'')]$ should be the interval "straightened out" when one takes the *convex hull* of the cumulative virtual value function. Without the assumption on the monotonicity of $\varphi_i(\cdot)$ on $[0, a) \cup (b, \infty]$, we may have multiple regions to iron and they all correspond to the straightened out regions when one takes the convex hull for the plot of the "cumulative virtual value" against the cumulative probabilities. We therefore obtain the following:

**Definition 1** (Ironed virtual values.)**.** For a bidder with virtual value function $\varphi(v)$, let $\{[a_1, b_1], \ldots, (a_k, b_k)\}$ be the intervals that are straightened out when one takes the convex hull for the plot of the cumulative virtual value $\int_0^v f(v)\varphi_i(v)\,\mathrm{d}v$ plotted against the cumulative probability $F(v)$. The *ironed virtual value* $\tilde{\varphi}(v)$ is

$$\tilde{\varphi}(v) = \begin{cases} \frac{\int_{a_i}^{b_i} \varphi(t) f(t)\,\mathrm{d}t}{F(b_i) - F(a_i)}, & v \in [a_i, b_i]; \\ \varphi(v), & \text{otherwise.} \end{cases}$$

**Theorem 3.** *The revenue of any incentive compatible auction with allocation rules* $(x_i(\cdot))_i$ *has revenue at most* $\sum_i \mathbf{E}_{v_i \sim F_i}[x_i(v_i)\tilde{\varphi}_i(v_i)]$.

With an argument similar to the case of regular distributions (i.e., pointwise optimality immediately implies optimality in expectation), we have

**Corollary 1** (Optimal auctions for general value distributions (Myerson, 1981).)**.** *In a single item auction where bidders' values are drawn independently, a revenue optimal auction allocates to the bidder with the highest non-negative ironed virtual value. The resulting allocation rule is monotone and can be supported by a payment that guarantees incentive compatibility.*

**Remark 1.** The "distance" or "cumulative virtual value" provided valuable insight on the regions to iron. One may justifiably wonder what it is, that is, what economic meaning it has, if any. Recall the cumulative virtual value for value $v$ is $\int_0^v \varphi(t)f(t)\,\mathrm{d}t/F(v)$. If we apply Lemma 1 in a formal manner (and forgetting about the monotonicity constraint), this would be the revenue of a mechanism that allocates to a bidder only when her value is below $v$ and otherwise does not sell — of course there is no incentive compatible mechanism that implements such an allocation rule, but the "complement" of it, which sells only when the value is above $v$, is implementable, by simply a posted price at $v$. Therefore the cumulative virtual values are the "reverse" of the revenues obtainable by posted prices. In Section 3 we rederive Myerson's theorem starting from the so-called revenue curves.

## 3 Derivation of Myerson's Auction Using Revenue Curves

This section sketches the key steps in the derivation of Myerson's mechanism, including ironing, using revenue curves.[3] This alternative perspective has the advantage that it gives Myerson's revenue optimal auction an interpretation that is analogous to a monopolistic pricing problem (first observed by Bulow and Roberts (1989)), and sheds light on many single-item pricing problems.

We derive Myerson (1981)'s optimal mechanism in six steps. Seeing (ironed) virtual values as the derivatives of the (ironed) revenue curve comes from Bulow and Roberts (1989), although the actual proof idea here comes from Alaei et al. (2013).

1. **Characterization of Bayesian incentive compatible mechanisms.** Every BIC mechanism has monotone allocation rule, i.e., $x_i(v_i)$ is nondecreasing with $v_i$. Moreover, the expected payment is determined by the allocation rule: $p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(s)\,\mathrm{d}s$.

   **Note:** The characterzation is really more about IC than about BIC. For example, any DSIC mechanism must have its allocation rule $x_i(v_i, v_{-i})$ monotone in $v_i$ given any $v_{-i}$, and the payment $p_i(v_i, v_{-i})$ is also determined as $v_i x_i(v_i, v_{-i}) - \int_0^{v_i} x_i(s, v_{-i})\,\mathrm{d}s$.

2. **Decomposition into step functions.** Any monotone allocation rule is the convex decomposition of step functions. In other words, the function $x_i(v_i)$ can be written as the weighted sum of some step functions, and the weights are nonnegative and sum to 1. (In the continuous case, we have an integral instead of a sum.) We would like to determine the weights or density of these step functions in this decomposition. Let us try to determine the weight of

[3]For full details, the reader is referred to the textbook by Jason Hartline. Later versions of this note may fill in more details.

the step function that jumps from 0 to 1 at $v$. Intuitively, the value $v$'s allocation is $x_i'(v)\,\mathrm{d}v$ more than the value that is slightly below it; this difference should be the probability that this particular step function is used, and so the weight should be $x_i'(v)\,\mathrm{d}v$.

3. **Calculating revenue using posted prices.** By revenue equivalence (i.e., that payment is determined by the allocation rule), to implement any monotone allocation rule, it is equivalent to randomize over a set of allocation functions that are step functions, where the probability of running the step function that jumps from 0 to 1 at value $v$ is $x_i'(v)\mathrm{d}v$. Such a step function is implemented by a posted price at $v$, and its expected revenue is $v(1-F_i(v))$. The expected revenue of any allocation rule $x_i$ is therefore

$$\int_0^\infty [v(1-F(v))]x_i'(v)\,\mathrm{d}v. \tag{1}$$

   Note that the integral over $v$ here is not with respect to the density $f_i$. We can already do an integral by part at this point, and using the fact that $v(1-F(v))$ evaluates to 0 at both 0 and $\infty$, this integral is equal to

$$\int_0^\infty x(v)[vf_i(v)-(1-F_i(v))]\,\mathrm{d}v = \int_0^\infty x_i(v)\left[v-\frac{1-F_i(v)}{f_i(v)}\right]f_i(v)\,\mathrm{d}v.$$

   The last step, extracting the factor $f_i(v)$ from the bracket, gives us the expression for *virtual surplus* with respect to the *virtual value*. This is the expression in Myerson's original proof. Note that here, by having the density function as the measure, it is as if we are taking expectation with respect to $v$ drawn from its original distribution. This meaning was not there in (1). This change of meaning in the integral variable is crucial, and it is one of the motivations for us to move from the value space to quantile space.

4. **Passing to the quantile space.** As we have seen, the integral (1) is with respect to a distribution of step functions (or posted prices), given by $x'(v)\,\mathrm{d}v$, where the measure for $v$ itself is uniform. We could as well carry over to a uniform distribution on the compact domain $[0,1]$ through the mapping $\psi_i(v)=1-F_i(v)$. $\psi_i(v)$ is called the *quantile* of the value $v$. Let $y_i:[0,1]\to[0,1]$ be the quantile allocation function, that is, $y_i(q)=x_i(\psi_i^{-1}(q))$. In the decomposition of the allocation rule, the step function jumping at $v$ then has weight/density $-y_i'(q)\,\mathrm{d}q$ evaluated at $q=\psi_i(v)$.[4] Let $R_i(q)=q\cdot F_i^{-1}(1-q)$ be the revenue of the step function at (or, equivalently, the posted price of) $v=\psi^{-1}(q)$, the revenue (1) can be rewritten, in terms of quantiles, as

$$\int_0^1 R_i(q)(-y_i'(q))\,\mathrm{d}q = \int_0^1 R_i'(q)y_i(q)\,\mathrm{d}q, \tag{2}$$

   where the equality again follows by integral by part. $R_i(q)$ is called the *revenue curve.*

   **Definition 2.** A distribution $F_i$ is said to be regular if its regular curve is concave.

[4] As a sanity check, $-y_i'(q) = -\frac{\mathrm{d}x_i(\psi_i(v))}{\mathrm{d}v}\cdot\frac{\mathrm{d}v}{\mathrm{d}\psi_i(v)}\Big|_{v=\psi_i^{-1}(q)} = x_i'(\psi_i^{-1}(q))$, agreeing with our calculation before.

**Remark:** Passing to the quantile space may seem strange at first. One of the advantages of this switch is that it facilitates a perspective change. Instead of thinking about $v(1 - F(v))$, the revenue of a certain posted price, $R(q)$ suggests the revenue of a selling strategy that sells with *ex ante* probability $q$: setting a price at $\psi_i^{-1}(q)$ is only one of such strategies. This perspective immediately leads to a more general definition of revenue curve and ironing itself.

5. **Generalization of revenue curves and Ironing.** Let $\tilde{R}_i(q)$ be the optimal revenue extractable from bidder $i$ with an incentive compatible mechanism that sells with ex ante probability $q$. Then obviously $\tilde{R}_i(q) \geq R_i(q)$ for any $q \in [0, 1]$. Furthermore, by step 2, any IC mechanism itself can be implemented by a distribution over posted prices. Therefore $\tilde{R}_i(q)$ is simply the concave hull of $R_i(q)$.[5]

   **Note:** For any $q$ where $\tilde{R}_i(q) > R_i(q)$, the revenue of the posted price $\psi_i^{-1}(q)$ is less than the revenue of a distribution over two other posted prices, whose expected ex ante selling probability is just $q$.

   Now, the revenue of any BIC mechanism from bidder $i$ is

$$\int_0^1 R_i'(q) y_i(q)\,\mathrm{d}q = \int_0^1 R_i(q)(-y_i'(q))\,\mathrm{d}q \leq \int_0^1 \tilde{R}_i(q)(-y_i'(q))\,\mathrm{d}q = \int_0^1 \tilde{R}_i'(q) y_i(q)\,\mathrm{d}q. \quad (3)$$

6. **The optimal mechanism.** The optimal mechanism maximizes its revenue with respect to the RHS of (3), and in fact achieves it. By the inequality in (3), such a mechanism maximizes the revenue as well, with equality therein attained.

   Recall that $y_i(q)$ is the allocation of bidder $i$ when her value is $v = \psi_i^{-1}(q)$. Therefore, in order to maximize $\int_0^1 \tilde{R}_i'(q) y_i(q)\,\mathrm{d}q$, the optimal mechanism solicits bids $v_1, \ldots, v_n$, and maps them to quantiles $\psi_1(v_1), \ldots, \psi_n(v_n)$, then observes the corresponding $\tilde{R}_1'(\psi_1(v_1)), \cdots, \tilde{R}_n'(\psi_n(v_n))$. If the maximum among these is above zero, then allocate the item to this bidder; otherwise, do not sell.

   **Remark:** The quantity $\tilde{R}_i'(q)$ is the *ironed virtual value* of bidder $i$'s type that has quantile $q$. It is "ironed" because in any region $(q_1, q_2)$ where $\tilde{R}_i$ is strictly greater than $R_i$, $\tilde{R}_i$ is a straight line and has all types in that region have the same ironed virtual value, and therefore, in the optimal mechanism, they are all treated the same. Equivalently, any posted price whose selling probability lies in $(q_1, q_2)$ is used with probability 0 in the optimal mechanism. In other words, the allocation rule is flat on $(q_1, q_2)$. This is also necessary for the equality in (3) to be attained.

## References


Alaei, S., Fu, H., Haghpanah, N., and Hartline, J. D. (2013). The simple economics of approximately optimal auctions. In *54th Annual IEEE Symposium on Foundations of Computer Science, Berkeley, CA, USA*, FOCS'13, pages 628–637.

Bulow, J. and Roberts, J. (1989). The simple economics of optimal auctions. *Journal of Political Economy*, 97(5):1060–90.


[5] Equivalently, the region enclosed by $\tilde{R}_i(q)$ with the $q$-axis is the convex hull of that enclosed by $R_i(q)$ and the $q$-axis.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73.