

# On Optimization of Tax Functions

T. V. Bogachev<sup>1\*</sup> and S. N. Popova<sup>2,1,3\*\*</sup>

<sup>1</sup> National Research University Higher School of Economics, Moscow, 101000 Russia

<sup>2</sup> Moscow Institute of Physics and Technology (State University), Dolgoprudnyi,  
Moscow Oblast, 141701 Russia

<sup>3</sup> Lomonosov Moscow State University, Moscow, 101000 Russia

Received January 10, 2020; in final form, January 10, 2020; accepted February 10, 2020

**Abstract**—We consider the problem of maximization of an integral functional on the space of increasing functions, which is motivated by economic concerns for tax mechanisms optimization. An analytical description of the optimal value is obtained, as well as an approximation method for finding the solution.

**DOI:** 10.1134/S000143462101020X

Keywords: *nonlinear optimization, optimal taxation.*

## 1. INTRODUCTION

Optimization of taxation mechanisms has become a subject of pure mathematical research only comparatively recently. This is natural, because such systems contain a very large number of parameters, especially when considered in detail. In order to overcome this complexity, researchers usually introduce a simple model of an economic environment (we will describe it later) and address the issue within it. For example, the authors of the recent papers [1] and [2] use dynamical methods for analyzing tax function properties. After making some assumptions, it is possible to find an analytical solution to the optimization problem in that simplified environment, which is the main result of the present paper.

Several modifications of the optimization problem have been explored in many papers in mathematical economics; we only mention the celebrated paper [3] and the survey [4] by the Nobel prize winner Mirrlees and also the papers [5] and [6]. The exact problem considered in this paper has been motivated by the papers [1] and [2] and our discussions with A. Tsyvinski. This problem is concerned with the maximization of an integral functional on the space of increasing functions subject to an additional nonlinear constraint, which generates quite singular objects. This is why a rigorous mathematical analysis of the problem leads to interesting questions in the theory of functions, in contrast to most works in applied economics, where all functions are assumed to be differentiable and to have zero derivatives at all extremum points. It should also be noted that this problem differs from the classical problems of optimal control (see, e.g., [7][8]).

## 2. STATEMENT OF THE PROBLEM

Assume that every economic agent is an object of a certain class completely characterized by its productivity type  $\theta \in \Theta \subset \mathbb{R}_+$ . Each agent's activity is characterized by labor effort  $l > 0$ . As a result, we have the agent's individual income  $u = \theta l$ . This income is liable to tax  $T(u)$ , which does not exceed the income and does not increase.

We also introduce the notion of the costs uncurred by an agent exerting effort  $l$ ; it is described by an increasing function  $f(l)$ . In this paper, we make a change of variables and consider the problem for variables  $u$  and  $\theta$  instead of  $\theta$  and  $l$ , which transforms  $f(l)$  into  $f(u, \theta)$ . We impose the following constraints on the functions mentioned above:

---

\*E-mail: tvbogach@gmail.com

\*\*E-mail: popovaclaire@mail.ru

- $f(u, \theta)$  is a nonnegative continuous function defined for  $u \geq 0$  and  $\theta > 0$  (or for  $\theta$  belonging to some interval  $[\theta_{\min}, \theta_{\max}] \subset (0, +\infty)$ );
- $f(0, \theta) = 0$  for every  $\theta$ ; the function  $f(u, \theta)$  increases in  $u$ ;
- the derivative  $\partial_\theta f(u, \theta)$  exists and is continuous in  $\theta$  and is decreasing in  $u$ ; there exists a derivative  $\partial_u f(u, \theta) > 0$ ;
- $T: [0, +\infty) \rightarrow [0, +\infty)$  is a continuous increasing function,  $T(0) = 0$ , and the function given by  $u \mapsto u - T(u)$  is increasing.

We denote the class of such functions by  $\mathcal{T}$ .

Under these conditions,  $\partial_\theta f(u, \theta) \leq \partial_\theta f(0, \theta) = 0$  for all  $u$  and  $\theta$ ; so  $f(u, \theta)$  is decreasing in  $\theta$ .

Suppose also that, for every fixed  $\theta > 0$  and all  $u$  large enough, we have

$$u - f(u, \theta) < 0.$$

A typical example is the function  $f(u, \theta) = u^2/(2\theta^2)$ . For this function, the equality  $\partial_\theta f(0, \theta) = -u^2/\theta^3$  holds.

The class  $\mathcal{T}$  can be described as the set of all increasing 1-Lipschitz functions vanishing at zero. Indeed, from the increasing of  $T(u)$  and  $u - T(u)$  we obtain

$$0 \leq T(u+h) - T(u) \leq u+h-u = h \quad \text{for all } h \geq 0.$$

Note that  $T(u) \leq u$ . For example, the class  $\mathcal{T}$  contains all linear functions  $ku$  with  $0 \leq k \leq 1$ .

For a given type  $\theta$  and a fixed tax function  $T$ , we denote

$$M(T, \theta) = \max_{u \geq 0} (u - T(u) - f(u, \theta)), \quad (2.1)$$

which corresponds to the maximal net income of the agent after paying taxes and other costs. It has the meaning of the optimal effort of an agent of the given type in the given tax system. Let  $u_T(\theta)$  denote the minimal point of maximum for (2.1). Such a point exists due to the constraints imposed on  $f$  and  $T$ . We set  $u_T(0) = 0$ .

Suppose that we are given a Borel probability measure  $P = p dx$  on  $(0, +\infty)$ . We will assume that either the density  $p$  is positive on  $(0, +\infty)$  or the measure  $P$  is concentrated on a bounded interval and the density  $p$  is positive on this interval. We set

$$F(\theta) = P([\theta, +\infty)) = \int_\theta^{+\infty} p(t) dt.$$

We will assume that there is a  $P$ -integrable locally bounded function  $R > 0$  for which  $u - f(u, \theta) \leq 0$  for all  $u \geq R(\theta)$ . Then  $u_T(\theta) \leq R(\theta)$ , so the estimate  $T(u_T(\theta)) \leq u_T(\theta) \leq R(\theta)$  implies that the functions  $T(u_T(\theta))$  and  $u_T(\theta)$  are integrable with respect to  $P$ .

Then we define the government revenue

$$R(T) = \int_{\Theta} T(u_T(\theta)) P(d\theta), \quad (2.2)$$

where we adhere to a scenario assuming that each agent exerts optimal effort and strives for increasing net income.

We study the problem of finding the value

$$J(f, P) = \sup_{T \in \mathcal{T}} R(T), \quad (2.3)$$

that is, of maximizing the collected tax.

Here it is of interest to find both the maximal value and the maximizing mapping  $T$ . Our main results give a constructive answer to the first question for a broad class of functions  $f$ , and for more special functions, we manage to obtain some information about the optimal mapping  $T$ .

We observe that it follows from what has been said above that  $J(f, P) \leq \|R\|_{L^1(P)}$ .

Note that, although we have to maximize the integral of the composition of  $T$ , simply increasing this function  $T$  may decrease the integral, because the point of maximum  $u_T$  may become smaller. For example, when taxation is excessively high, the actually collected tax may be zero.

3. MAIN RESULTS

Let  $\mathcal{U}$  denote the set of all increasing left continuous functions  $u: [0, +\infty) \rightarrow [0, +\infty)$  such that  $u(0) = 0$  and, for all  $\theta > 0$ , the inequality

$$u(\theta) \leq R(\theta), \quad \partial_u f(u(\theta), \theta) \leq 1, \quad f(u(\theta), \theta) \leq u(\theta)$$

holds. For every function  $u \in \mathcal{U}$ , we define the right inverse function by the formula

$$v(s) = \sup\{t : u(t) \leq s\}.$$

The function  $v$  also increases and is left continuous. In addition,  $u(v(s)) = s$  at the points in the range of  $u$ , since, for such points,  $v(s)$  is the maximal point in  $u^{-1}(s)$ .

The main result of the paper is the following theorem.

**Theorem 1.** *The following equalities hold:*

$$\begin{aligned} J(f, P) &= \sup_{u \in \mathcal{U}} \int \left( u(\theta) - f(u(\theta), \theta) + \int_0^\theta \partial_\theta f(u(\tau), \tau) d\tau \right) P(d\theta) \\ &= \sup_{u \in \mathcal{U}} \int [u(\theta) - f(u(\theta), \theta)] P(d\theta) + \int \partial_\theta f(u(\theta), \theta) F(\theta) d\theta, \end{aligned} \tag{3.1}$$

where the supremum can be taken over increasing infinitely differentiable functions  $u$  with  $u' > 0$ . Approximations for the supremum in  $T$  can be obtained by means of mappings of the form

$$T(s) = s - f(s, v(s)) + \int_0^{v(s)} \partial_\theta f(u(\tau), \tau) d\tau,$$

where  $u \in \mathcal{U}$  is an infinitely differentiable function and  $v$  is the inverse function for  $u$ .

**Proof.** Given a fixed mapping  $T \in \mathcal{T}$ , we denote the function  $u_T(\cdot)$  simply by  $u(\cdot)$  for brevity. Let us show that the function  $\theta \mapsto u(\theta)$  possesses the following properties:

- (1)  $u \in \mathcal{U}$ ;
- (2) whenever  $\theta > 0$  and  $0 \leq u < u(\theta)$ , the inequality

$$u - f(u, \theta) < u(\theta) - f(u(\theta), \theta)$$

holds.

We first verify that the function  $\theta \mapsto u(\theta)$  increases. Let  $\theta_1 < \theta_2$ . Then

$$u(\theta_1) - T(u(\theta_1)) - f(u(\theta_1), \theta_1) > u - T(u) - f(u, \theta_1)$$

whenever  $u < u(\theta_1)$ . In addition,  $f(u, \theta_1) - f(u, \theta_2)$  increases in  $u$  if  $\theta_1 < \theta_2$  by the condition that the function  $-\partial_\theta f(u, \theta)$  increases in  $u$ . Therefore,

$$u(\theta_1) - T(u(\theta_1)) - f(u(\theta_1), \theta_2) > u - T(u) - f(u, \theta_2)$$

for all  $u < u(\theta_1)$ , i.e.,  $u(\theta_2) \geq u(\theta_1)$ .

We now verify that this function is left continuous. Suppose that  $\lim_{\theta \rightarrow \theta_0 - 0} u(\theta) = u_0 < u(\theta_0)$  for some  $\theta_0 > 0$ . For every  $\theta < \theta_0$  we have

$$u(\theta) - T(u(\theta)) - f(u(\theta), \theta) \geq u(\theta_0) - T(u(\theta_0)) - f(u(\theta_0), \theta),$$

since  $u(\theta)$  is a point of maximum of (2.1). Passing to the limit as  $\theta$  tends to  $\theta_0$  from the left, we obtain

$$u_0 - T(u_0) - f(u_0, \theta_0) \geq u(\theta_0) - T(u(\theta_0)) - f(u(\theta_0), \theta_0).$$

This means that, actually, this is an equality, since  $u(\theta_0)$  is a point of maximum. Hence  $u_0$  is also a point of maximum, which contradicts the fact that  $u(\theta_0)$  is the minimal point of maximum of (2.1). Thus, the function  $u(\cdot)$  is left continuous. By definition,  $0 \leq u_T \leq R(\theta)$  and  $u_T(0) = 0$ .

Finally, we show that property (2) is fulfilled. Let  $\theta > 0$  and  $0 \leq u < u(\theta)$ . Then

$$u - T(u) - f(u, \theta) < u(\theta) - T(u(\theta)) - f(u(\theta), \theta),$$

because  $u(\theta)$  is the minimal point of maximum of (2.1). It follows that

$$u - f(u, \theta) < u(\theta) - f(u(\theta), \theta)$$

since  $T$  is increasing. Thus, property (2) is established. It obviously yields the estimate  $\partial_u f(u(\theta), \theta) \leq 1$ , which completes the proof of the inclusion  $u \in \mathcal{U}$ .

Since  $\max_{u \geq 0} (u - T(u) - f(u, \theta)) \geq 0$ , we have

$$T(u(\theta)) \leq u(\theta) - f(u(\theta), \theta) \quad \text{for all } \theta > 0.$$

Therefore, we have

$$T(u(\theta)) = u(\theta) - f(u(\theta), \theta) - r(\theta), \quad \text{where } r(\theta) \geq 0.$$

Let us verify that the function  $r$  is increasing. Let  $\theta_1 < \theta_2$ . Then  $u(\theta_1) \leq u(\theta_2)$  and

$$u(\theta_1) - T(u(\theta_1)) - f(u(\theta_1), \theta_2) \leq u(\theta_2) - T(u(\theta_2)) - f(u(\theta_2), \theta_2),$$

because  $u(\theta_2)$  is a point of maximum of (2.1). It follows that

$$r(\theta_2) \geq r(\theta_1) + f(u(\theta_1), \theta_1) - f(u(\theta_1), \theta_2).$$

In particular,  $r(\theta_2) \geq r(\theta_1)$ , since  $f(u(\theta_1), \theta_1) \geq f(u(\theta_1), \theta_2)$ .

However, we have proved even more, namely, the estimate

$$\frac{r(\theta_2) - r(\theta_1)}{\theta_2 - \theta_1} \geq \frac{f(u(\theta_1), \theta_1) - f(u(\theta_1), \theta_2)}{\theta_2 - \theta_1}.$$

Let us show that this estimate implies the inequality

$$Dr \geq -\partial_\theta f(u(\theta), \theta)$$

in the sense of distributions (see, e.g., [9]), where  $Dr$  does not denote the pointwise derivative, but stands for the measure that is the generalized derivative of the function  $r$  (this measure is finite on all intervals  $[0, N]$ ). We first observe that the nonnegative increasing function  $r$  has a finite limit  $r_0 = \lim_{\theta \rightarrow 0+} r(\theta)$ . In addition, the increasing function  $r$  has the usual derivative  $r'$  almost everywhere, and  $r'$  is integrable on all intervals  $[0, N]$ , since  $r_0$  is finite. It follows from this and the estimate  $r'(\theta) \geq -\partial_\theta f(u(\theta), \theta) \geq 0$  that the function  $\partial_\theta f(u(\theta), \theta)$  is integrable in  $\theta$  on all intervals  $[0, N]$ . We observe that the function  $\partial_\theta f(u, \theta)$  with fixed  $u$  may fail to be integrable in  $\theta$  up to zero: take  $f(u, \theta) = u^2/\theta^2$ .

From the estimate proved above for the difference quotient we obtain

$$n \left[ r \left( \theta + \frac{1}{n} \right) - r(\theta) \right] \geq n \left[ f(u(\theta), \theta) - f \left( u(\theta), \theta + \frac{1}{n} \right) \right].$$

Hence, for every smooth function  $g \geq 0$  with support in  $(0, +\infty)$ , we have the inequality

$$\int n \left[ r \left( \theta + \frac{1}{n} \right) - r(\theta) \right] g(\theta) d\theta \geq \int n \left[ f(u(\theta), \theta) - f \left( u(\theta), \theta + \frac{1}{n} \right) \right] g(\theta) d\theta.$$

Making a change of variable, we can write the left-hand side in the form

$$\int nr(s) \left[ g \left( s - \frac{1}{n} \right) - g(s) \right] ds \rightarrow - \int r(s) g'(s) ds$$

as  $n \rightarrow \infty$ . On the right-hand side of the indicated inequality, under the integral sign, we have a nonnegative function whose limit as  $n \rightarrow \infty$  equals  $-g(\theta) \partial_\theta f(u(\theta), \theta)$ . By Fatou's theorem,

$$- \int g(\theta) \partial_\theta f(u(\theta), \theta) d\theta \leq - \int r(s) g'(s) ds,$$

where the right-hand side coincides with the action of the generalized function (distribution)  $Dr$  on  $g$ , i.e., equals the integral of  $g$  against the measure  $Dr$ .

Thus, the desired inequality in the sense of distributions is established. This inequality, along with the integrability of the function  $\theta \mapsto \partial_\theta f(u(\theta), \theta)$  on all intervals  $[0, N]$ , implies the inequality

$$r(\theta) \geq r(\theta) - r_0 \geq - \int_0^\theta \partial_\theta f(u(\tau), \tau) d\tau.$$

Thus, we have

$$T(u(\theta)) \leq u(\theta) - f(u(\theta), \theta) + \int_0^\theta \partial_\theta f(u(\tau), \tau) d\tau.$$

Therefore, the left-hand side of (3.1) is not greater than the right-hand side. Note that the conditions  $f \geq 0$ ,  $\partial_\theta f \leq 0$ , and  $T \geq 0$  yield that, on the right-hand side of the last estimate, every term does not exceed  $u(\theta)$  in absolute value.

We now fix a function  $u \in \mathcal{U}$ . Let us take the sequence of infinitely differentiable functions  $u_n \geq 0$  on  $[0, +\infty)$  defined by the formula

$$u_n(\theta) = \int_0^\infty u\left(\theta - \frac{s}{n}\right) \varrho(s) ds,$$

where  $\varrho$  is an infinitely differentiable probability density vanishing on  $(-\infty, 0]$  and positive on  $(0, +\infty)$ . We observe that

$$0 \leq u_n(\theta) \leq u(\theta), \quad u'_n(\theta) = n \int_0^\theta \varrho(n(\theta - t)) Du(dt).$$

In addition, the values  $u_n(\varphi)$  increase to  $u(\varphi)$  for all  $\varphi \geq 0$ . We can assume that  $u$  is not identically zero. Then there exists an interval  $[0, a]$  on which the nonnegative measure  $Du$  is not zero. Let  $a_0$  be the infimum of such  $a$ . It is seen from the previous formula that  $u'_n(t) > 0$  for all  $t > a_0$ . If  $a_0 > 0$ , then  $u(\theta) = 0$  for all  $\theta \leq a_0$  and  $u(\theta) > 0$  for all  $\theta > a_0$ . Then the functions  $v_n = u_n^{-1}$  exist and are smooth, and  $v_n(0) = a_0$ . Let us now consider the smooth functions

$$T_n(s) = s - f(s, v_n(s)) + \int_0^{v_n(s)} \partial_\theta f(u_n(\tau), \tau) d\tau.$$

We observe that

$$\begin{aligned} T'_n(s) &= 1 - \partial_u f(s, v_n(s)) - \partial_\theta f(s, v_n(s))v'_n(s) + \partial_\theta f(s, v_n(s))v'_n(s) \\ &= 1 - \partial_u f(s, v_n(s)) \geq 0, \end{aligned}$$

since  $v_n(s) \geq v(s)$  in view of the estimate  $u_n \leq u$ , which ensures the inequality

$$\partial_u f(s, v_n(s)) \leq \partial_u f(s, v(s)) = \partial_u f(u(v(s)), v(s)) \leq 1.$$

Since  $\partial_u f \geq 0$ , we also have  $T'_n(s) \leq 1$ . In addition,  $T_n(0) = 0$ , because, for any  $\tau \leq a_0$ , we have  $u_n(\tau) = 0$ , whence  $\partial_u f(u_n(\tau), \tau) = 0$ . Thus,  $T_n \in \mathcal{T}$ .

Let us show that, for the mapping  $T_n$ , the function  $u_n$  is exactly  $u_{T_n}$ . Indeed,

$$u - T_n(u) - f(u, \theta) = f(u, v_n(u)) - f(u, \theta) - \int_0^{v_n(u)} \partial_\theta f(u_n(\tau), \tau) d\tau.$$

The derivative of this expression equals

$$\partial_u f(u, v_n(u)) - \partial_u f(u, \theta).$$

By our condition on  $\partial_u f$ , the equality to zero is only possible for  $\theta = v_n(u)$ , that is, for  $u = u_n(\theta)$ .

Let us verify that the integrals of the functions  $T_n(u_n(\theta))$  against the measure  $P$  tend to the integral of  $T(u(\theta))$ . We have

$$T_n(u_n(\theta)) = u_n(\theta) - f(u_n(\theta), \theta) + \int_0^\theta \partial_\theta f(u_n(\tau), \tau) d\tau.$$

Since  $u_n(\theta)$  increases to  $u(\theta)$ , the convergence of the integrals holds for the first two terms. The negative functions  $\partial_\theta f(u_n(\tau), \tau)$  decrease to  $\partial_\theta f(u(\tau), \tau)$ , which implies the decreasing of their integrals over the interval  $[0, \theta]$  to the integral of  $\partial_\theta f(u(\tau), \tau)$  for every  $\theta$ . Since  $f \geq 0$ ,  $\partial_\theta f \leq 0$ , and  $T_n \geq 0$ , we have the estimates

$$f(u_n(\theta), \theta) \leq f(u(\theta), \theta) \leq u(\theta),$$

$$\int_0^\theta |\partial_\theta f(u_n(\tau), \tau)| d\tau \leq \int_0^\theta |\partial_\theta f(u(\tau), \tau)| d\tau \leq u_n(\theta) \leq u(\theta).$$

Taking into account the integrability of  $u(\theta)$  with respect to  $P$ , we obtain the desired convergence of the integrals with respect to the measure  $P$ .  $\square$

#### 4. CONCLUSIONS

So far, we have reduced one supremum over a class of functions to another supremum over another class of functions, but without evaluating intermediate maxima, which simplifies the problem, although does not enable us to find the solution explicitly. However, we now introduce a number of conditions on  $f$  and  $p$ , which lead to further simplifications.

For every  $\theta > 0$  (or for every  $\theta$  from a given interval on which the measure  $P$  is concentrated), we consider the quantity

$$M(\theta) = \max_{u \geq 0} [up(\theta) - f(u, \theta)p(\theta) + \partial_\theta f(u, \theta)F(\theta)].$$

Due to our conditions on  $f$ , this maximum exists. Suppose that there is a function  $u(\cdot) \in \mathcal{U}$  for which this maximum is attained at some point  $u(\theta)$  for every  $\theta > 0$ . Then, obviously,

$$J(f, P) = \int_0^\infty M(\theta) d\theta = \int \max_{u \geq 0} [up(\theta) - f(u, \theta)p(\theta) + \partial_\theta f(u, \theta)F(\theta)] d\theta.$$

If this function  $u(\cdot)$  is continuously differentiable and has positive derivative, then, as in the proof of the theorem, by means of its inverse  $v$ , we can construct a function  $T \in \mathcal{T}$ , for which  $u = u_T$ , so  $T$  gives a maximum in our problem. However, one can hardly evaluate  $T$  in nontrivial examples.

Let us consider the special case

$$f(u, \theta) = \frac{u^2}{2\theta^2}. \quad (4.1)$$

In this case,  $M(\theta)$  is attained at the point

$$u(\theta) = \frac{p(\theta)\theta^3}{p(\theta)\theta + 2F(\theta)}. \quad (4.2)$$

If this function is continuous and increasing, then a straightforward verification shows that it belongs to the class  $\mathcal{U}$ . Therefore,

$$J(f, P) = \frac{1}{2} \int \frac{p^2(\theta)\theta^3}{p(\theta)\theta + 2F(\theta)} d\theta. \quad (4.3)$$

For a differentiable function  $p$  we arrive at the condition

$$\frac{2p(\theta)}{F(\theta)} + \frac{2}{\theta} + \frac{p'(\theta)}{p(\theta)} \geq 0.$$

This condition is fulfilled for the uniform distributions in intervals (and, more generally, for distributions in interval with nondecreasing densities), for the standard Gaussian measure, for the exponential distribution, and for the Cauchy distribution.

With the aid of the explicitly found function  $u$ , it is possible in principle to find the corresponding functions  $v$  and  $T$ , but one should not expect to have simple explicit formulas for them. For example, if  $P$

is the uniform distribution in  $[0, 1]$ , then  $p = 1$  in  $[0, 1]$ ,  $F(\theta) = 2 - 2\theta$ , and  $u(\theta) = \theta^3/(2 - \theta)$ . In this case,  $v$  is found from the cubic equation

$$s(2 - v(s)) = v(s)^3.$$

Note that if we take only the linear functions  $T(u) = ku$  for  $f(u, \theta) = u^2/(2\theta^2)$ , then the maximum is attained for  $k = 1/2$  and equals

$$\frac{1}{4} \int \theta^2 p(\theta) d\theta.$$

This quantity is less than the exact maximum indicated above. On the other hand, it follows from the Cauchy inequality that we can obtain more than

$$\frac{1}{2} \int \theta^2 p(\theta) d\theta$$

for no  $T$ . Indeed,

$$u(\theta) - T(u(\theta)) - \frac{u(\theta)^2}{2\theta^2} \geq 0,$$

whence

$$T(u(\theta)) \leq u(\theta) - \frac{u(\theta)^2}{2\theta^2} = -\frac{1}{2} \left( \frac{u(\theta)}{\theta} - \theta \right)^2 + \frac{1}{2} \theta^2 \leq \frac{1}{2} \theta^2.$$

Thus, the obtained value is between the quarter and the half of the second moment of  $p$ .

Let us consider a more general example. Suppose, in addition, that

$$\partial_u^2 f(u, \theta) > 0, \quad \partial_u \partial_\theta^2 f(u, \theta) \geq 0, \quad \partial_u^2 \partial_\theta f(u, \theta) \leq 0.$$

If the measure  $P$  is concentrated on an interval, then, under these conditions,  $\theta$  is taken from the corresponding interval. We set

$$G(\theta) = \frac{F(\theta)}{p(\theta)}.$$

Let us show that if  $G'(\theta) \leq 1$  for all  $\theta > 0$ , then there is a unique point  $u(\theta)$  at which  $M(\theta)$  is attained, and  $M(\theta)$  increases in  $\theta$ .

It follows from the indicated conditions that the function

$$p(\theta) - \partial_u f(u, \theta)p(\theta) + \partial_u \partial_\theta f(u, \theta)F(\theta)$$

is strictly decreasing (since it has negative derivative). Hence it can have only one zero. Let  $\theta_1 < \theta_2$ . Then

$$u - f(u, \theta_1) + \partial_\theta f(u, \theta_1)G(\theta_1) < u(\theta_1) - f(u(\theta_1), \theta_1) + \partial_\theta f(u(\theta_1), \theta_1)G(\theta_1)$$

for all  $u < u(\theta_1)$ , because  $u(\theta_1)$  is a point of maximum. Let us show that

$$f(u, \theta_1) - f(u, \theta_2) - \partial_\theta f(u, \theta_1)G(\theta_1) + \partial_\theta f(u, \theta_2)G(\theta_2)$$

increases in  $u$ ; it will follow from this that

$$u - f(u, \theta_2) + \partial_\theta f(u, \theta_2)G(\theta_2) < u(\theta_1) - f(u(\theta_1), \theta_2) + \partial_\theta f(u(\theta_1), \theta_2)G(\theta_2)$$

for all  $u < u(\theta_1)$ , that is,  $u(\theta_2) \geq u(\theta_1)$ . The increase of this function in  $u$  follows from the decrease in  $u$  of the function

$$\partial_\theta f(u, \theta) - (\partial_\theta f(u, \theta)G(\theta))'_\theta = \partial_\theta f(u, \theta)(1 - G'(\theta)) - \partial_\theta^2 f(u, \theta)G(\theta).$$

The condition  $G' \leq 1$  can be written in the form

$$-p' \leq 2p^2 F.$$

It is fulfilled for any increasing density on an interval (but is not fulfilled for the typical decreasing densities on the half-line).

**Remark.** We can take a sequence of functions  $T_n \in \mathcal{T}$ , for which the integrals of  $T_n(u_{T_n}(\theta))$  with respect to the measure  $P$  increase to  $J(f, P)$ . Since these functions are 1-Lipschitz and  $T_n(0) = 0$ , we can pick a subsequence uniformly converging on bounded intervals to some function  $T \in \mathcal{T}$ . The increasing functions  $u_{T_n}$  are estimated by the function  $R$ ; hence we can pass to a further subsequence (numbered by the same indices) that converges pointwise to some increasing function  $u$ . The function  $u(\theta) - T(u(\theta)) - f(u(\theta), \theta)$  is the pointwise limit of the functions  $u_{T_n}(\theta) - T_n(u_{T_n}(\theta)) - f(u_{T_n}(\theta), \theta)$ . Indeed, it follows by the uniform Lipschitz continuity of the functions  $T_n$  that  $T_n(x_n) \rightarrow T(x)$  as  $x_n \rightarrow x$ . This implies that  $u(\theta)$  is a point of maximum of the function  $x - T(x) - f(x, \theta)$ . However, it is not clear whether this point of maximum is minimal. If this is true, then the limit mapping  $T$  gives the maximum of the function under consideration.

#### ACKNOWLEDGMENTS

The authors thank A. Tsyvinski for useful discussions.

#### FUNDING

This work was supported by the Russian Science Foundation under grant 17-11-01058 at Moscow State University.

#### REFERENCES

1. S. Steinerberger and A. Tsyvinski, *Tax Mechanisms and Gradient Flows*, [arXiv:1904.13276](https://arxiv.org/abs/1904.13276) (2019).
2. D. Sachs, A. Tsyvinski, and N. Werquin, “Nonlinear tax incidence and optimal taxation in general equilibrium,” *Econometrica* **88** (2), 469–493 (2020).
3. J. A. Mirrlees, “An exploration in the theory of optimum income taxation,” *Rev. Econom. Stud.* **38** (2), 175–208 (1971).
4. J. A. Mirrlees, “The theory of optimal taxation,” in *Handbooks in Econom.* III, Vol. 1: *Handbook of Mathematical Economics* (North-Holland, Amsterdam, 1986), pp. 1197–1249.
5. M. F. Hellwig, “Incentive problems with unidimensional hidden characteristics: a unified approach,” *Econometrica* **78** (4), 1201–1237 (2010).
6. E. Saez, “Using elasticities to derive optimal income tax rates,” *Rev. Econom. Stud.* **68**, 205–229 (2001).
7. V. M. Alekseev, V. M. Tikhomirov, and S. V. Fomin, *Optimal Control* (Consultants Bureau, New York, 1987).
8. S. M. Aseev and V. M. Veliov, “Another view of the maximum principle for infinite-horizon optimal control problems in economics,” *Russian Math. Surveys* **74** (6), 963–1011 (2019).
9. A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, Transl. from the 2nd Russian ed. (Dover, New York, 1975).