Concavifying the QuasiConcave (Published in a much shorter form that has mistakes fixed but is much harder to understand in *The Journal of Convex Analysis*, 24(4): 1239-62 (December 2017).)

August 14, 2014

Christopher Connell and Eric B. Rasmusen

Abstract

We revisit a classic question of Fenchel from 1953: Which quasiconcave functions can be concavified by a monotonic transformation? While many authors have given partial answers under various assumptions, we offer a complete characterization for all quasiconcave functions without a priori assumptions on regularity. In particular, we show that if and only if a real-valued function f is strictly quasiconcave, (except possibly for a flat interval at its maximum) and furthermore belongs to a certain explicitly determined regularity class, there exists a strictly monotonically increasing function g such that $g \circ f$ is strictly concave. Our primary new contribution is determining this precise minimum regularity class.

We prove this sharp characterization of quasiconcavity for continuous but possibly nondifferentiable functions whose domain is any Euclidean space or even any arbitrary geodesic metric space. Under the assumption of twice differentiability, we also establish simpler sufficient conditions for concavifiability on arbitrary Riemannian manifolds, which essentially generalizes those given by Kannai in 1977 for the Euclidean case. Lastly we present the approximation result that if a function f is either weakly or strongly quasiconcave then there exists an arbitrarily close strictly concavifiable approximation h to f.

Christopher Connell: Indiana University, Department of Mathematics, Rawles Hall, RH 351, Bloomington, Indiana, 47405. (812) 855-1883. Fax:(812) 855-0046. connell@indiana.edu.

Eric Rasmusen, Department of Business Economics and Public Policy, Kelley School of Business, Indiana University. BU 438, 1309 E. 10th Street, Bloomington, Indiana, 47405-1701. (812) 855-9219. Fax: 812-855-3354. erasmuse@indiana.edu, http://rasmusen.org.

This paper: http://rasmusen.org/papers/connell-rasmusen.quasiconcavity.pdf.
Mathematica and diagram files: http://rasmusen.org/papers/quasi.
Keywords: quasiconcavity, quasiconvexity, concavity, convexity, unique maximum, maximization. JEL Code: C61.

We would like to thank Per Hjertstrand and participants in the BEPP Brown Bag Lunch for comments, and Dan Zhao for research assistance.

1 Introduction

Quasiconcavity is a property of functions which, if strict, guarantees that a function defined on a compact set has a single, global, maximum and, if weak, a convex set of maxima. Most economic models involve maximization of a utility, profit, or social welfare function, and so assumptions which guarantee strict quasiconcavity are standard for tractability. Concavity is easier to understand than quasiconcavity, and concave functions on compact sets also have a single, global maximum, but concavity is a much stronger assumption.

In the rush of first-semester graduate economics, every economist encounters concavity and quasiconcavity and learns that they both have something to do with maximization. Why quasiconcavity has the word "concavity" in its name, however, is left unclear. He learns that the concepts are nested— that all concave functions are quasiconcave— but quasiconcave functions can look very different from concave functions. Figure 1 depicts functions that are strictly convex, strictly concave, and neither convex nor concave. The third function has no trace of the diminishing returns or negative second derivatives that we associate with concavity. But all three curves have a single maximum and are quasiconcave.

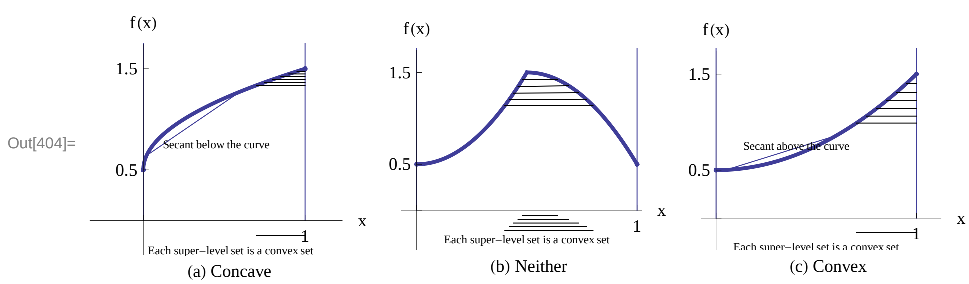

Figure 1: Three Strictly Quasiconcave Functions

[CHECK FIgure 1: I see some garbage to the left of the first diagram, and it looks too small.]

The purpose of this paper is to analyze a natural way to link concavity to quasiconcavity: by means of a monotonically increasing transformation. We will show that to say a sufficiently regular function is strictly quasiconcave is close to saying it can be made concave by a strictly increasing transformation.

Almost any sufficiently regular quasiconcave function can be concavified this way. Any function not quasiconcave cannot be.

The "almost" part of the conclusion must be qualified, but it can also be extended. The two qualifications are that after concavifying one monotonic portion, which can always be done even if the function is not differentiable, (i) its rate of change must not be "too near" zero or infinity, and (ii) it cannot have "too many" changes of its rate of change. (Both of these conditions will be made precise below.) An extension is that certain quasiconcave functions that are not only nondifferentiable but discontinuous can also be characterized this way, including all monotonic discontinuous functions.

Weakly quasiconcave functions which are not strictly quasiconcave away from the peak are never concavifiable. Nevertheless, weakly and strictly quasiconcave functions, even when not concavifiable, lie arbitrarily close to concavifiable functions. We show that the concavifiable functions are uniformly dense in the space of weakly quasiconcave functions.

Our topic goes back to the origins of the study of quasiconcavity, in DeFinetti (1949) and Fenchel (1953) (who invented the name, as Guerraggio & Molho (2004) explain in their history). The approach in our paper is most related to the economics literature exploring what sort of preferences can be represented by concave utility functions. Both DeFinetti (in his "second problem") and Fenchel investigated whether any quasiconcave function could be transformed into a concave function, which with related problems is surveyed in Rapcsak (2005) (see also Section 9 of Aumann (1975)). Kannai (1977, and 1981 more elaborately) treats the question in depth in the context of utility functions, giving conditions under which continuous convex preferences can be represented by concave utility functions. We will discuss Kannai's conditions in greater detail in connection with Theorem 4 below. Richter & Wong (2004) and Kannai (2005)s similarly address preferences over discrete sets.

The previous literature has asked when preference orderings that can be represented by differentiable functions can be represented by concave functions. We answer the complete problem of when quasiconcave functions can be concavified, whether they be utility, profit, social welfare, payoff, or inverse loss functions. We generalize from the previously analyzed differentiable functions in Euclidean space to strictly quasiconcave continuous functions on any non-flat manifold, which require a different approach. We also analyze the undiscussed cases of what may be done with weakly quasiconcave and discontinuous functions. In this way we hope to have shed new light on a classic problem in the fundamentals of optimization and to better understand the key concept of what kinds of functions have a single local maximum. We hope that this may eventually help by extending the use of convex optimization numerical techniques, though we do not explore them in the present paper.

We start the analysis with the easy case of a quasiconcave function defined over a finite domain of points. Next, we build up from strictly increasing differentiable functions $f\colon \mathbb{R}^1 \to \mathbb{R}$, to nonmonotonic differentiable functions, and then to nonmonotonic nondifferentiable continuous functions. The interesting necessary and sufficient conditions arise in the last two cases. We attain our greatest generality with functions on geodesic metric spaces in general, of which nondifferentiable functions $f\colon \mathbb{R}^n \to \mathbb{R}$ are a special case. We then backtrack to look at the interesting special case of twice-differentiable functions on manifolds, where sufficient conditions for concavifiability of a much simpler nature can be found. These latter conditions are closely related to those found in Kannai (1977), although very different in appearance and under slightly different assumptions.

Finally, we examine two classes of functions f to which our results do not apply without significant limitations: discontinuous functions and weakly quasiconcave functions. Our conclusion that a quasiconcave function is one that can be concavified does not apply to those two classes of functions generally. It does apply to a limited class of discontinuous functions, and a weakly quasiconcave function can always be approximated by a concavifiable one. We completely classify when discontinuous or weakly quasiconcave functions are concavifiable.

2 Formalization

Textbook expositions of quasiconcavity can be found in Kreps (1990 p. 67), Takayama (1985, p. 113), and Green, Mas-Colell and Whinston (1995, pp. 49, 943). Arrow & Enthoven (1961) is the classic article applying it to maximization Osborne (undated), Pogany (1999), and Wilson (2009) have useful notes on the Web. Definition 1 is one of several equivalent ways to define the term (Ginsberg (1973) and Pogany (1999) discuss variant definitions).

Definition 1. (QUASICONCAVITY) A function f defined on a subset $D \subset \mathbb{R}^n$ is *weakly quasiconcave* iff for any two distinct points $x', x'' \in D$ and any number $t \in (0,1)$ with $tx' + (1-t)x'' \in D$, we have

$$f(tx' + (1-t)x'') \geq \min\{f(x'), f(x'')\}. \tag{2.1}$$

Iff the inequality is strict whenever $x' \neq x''$, we say that f is *strictly quasiconcave.*

This definition does not require f to be differentiable, or even continuous. Often we assume that the domain D is convex, but we will use this generality for investigating finite domains as well.

An equivalent definition for weak quasiconcavity which we will exploit later on says that f is weakly quasiconcave if and only if its super-level sets are weakly convex; that is, the sets $S_c \equiv \{x \in D : f(x) \geq c\}$ are weakly convex for each c in the range of f as shown in Figure 1. For strict quasiconcavity, one requires strictly convex super-level sets together with the absence of horizontal line segments in the graph. (Figure 14, late in the paper, depicts a weakly but not strictly quasiconcave function.)

It is frequently plausible in economic applications that a function $f(x)$ being maximized is quasiconcave, which is convenient because quasiconcavity guarantees a unique supremum of $f(x)$ (which we will denote by $\boxed{f^*}$).[1] Strict quasiconcavity further guarantees a unique maximum on a closed set, $\boxed{m}$, where m: $f^* \equiv f(m)$. (If the quasiconcavity is only weak, there might be several x's such that $f^* = f(x)$, though the set of optimal x's will at least be convex.)

If $-f$ is strictly quasiconcave then f is *strictly quasiconvex.* If $-f$ is weakly quasiconcave then f is *weakly quasiconvex.*

[1]We will box definitions and examples where we think this useful for readers flipping back to find them.

Definition 2. (CONCAVITY) A function f defined on a subset $D \subset \mathbb{R}^n$ is *concave* iff for any two points $x', x'' \in D$ and any number $t \in (0,1)$ with $tx' + (1-t)x'' \in D$, we have

$$f(tx' + (1-t)x'') \geq tf(x') + (1-t)f(x''). \qquad (2.2)$$

Iff inequality (2.2) is strict, we say that f is strictly concave; otherwise it is weakly concave or simply "concave".

Figure 1(a) illustrates this definition, which says that the secant line must lie below the function. Every concave function is quasiconcave, but not every quasiconcave function is concave. That is because $\min(f(x'), f(x'')) \leq tf(x') + (1-t)f(x'')$. Quasiconcavity requires the function merely not to dip down and back up between x' and x'', but concavity requires it to rise faster than linear from the lower point to the upper one.

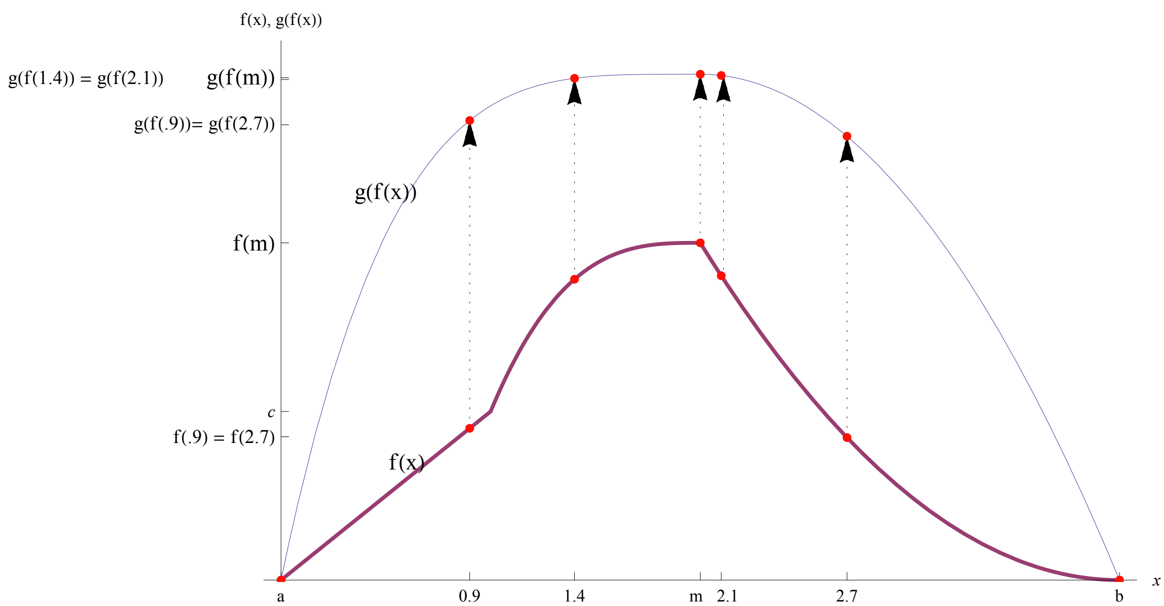

Figure 2: A Non-Concave Strictly Quasiconcave Function $f(x)$ that Can Be Concavified by a Monotonic $g(f)$

We will be looking at whether given a quasiconcave f we can always find a strictly monotonic function g that will transform f to a strictly concave $g \circ f$. Figure 2 shows an example of a strictly quasiconcave function $f(x)$ that is not concave, and a compound function $g(f(x))$ that is both strictly quasiconcave and strictly concave.[2]

In economics, we are interested in concavity and quasiconcavity because we usually assume diminishing returns to single activities as a result of mixtures being better than extremes (e.g., a taste for variety in food) or some activity levels being fixed (e.g., fixed capital in the short run). Thus, the set of utilities obtainable from a given budget will be a convex set, and the utility function will be concave in each good and quasiconcave over all of them. Figure 3 is an example in two dimensions that shows the function $f\colon \mathbb{R}^2_+ \to \mathbb{R}$ given by $f(x, y) = 50 \log(x) \log(y)$, which is strictly concave on $[e, \infty)^2 \subset \mathbb{R}^2_+$ but only quasiconcave on the larger domain $[1, \infty)^2 \subset \mathbb{R}^2_+$ despite the fact that each slice in any coordinate direction is concave on $[1, \infty]$. Note also that this function is monotone along all line segments, since the level sets are themselves line segments, which can only intersect any line segment at one point at most.

Most of the difficulties in concavifying quasiconcave functions arise even when the function's domain is one-dimensional Euclidean space, so $\mathbb{R}^1$ will be the focus of the first half of this article. Even $\mathbb{R}^2$ is much harder to visualize, as may be seen from Figure 4s example:

Figure 4's example: Fenchel's example. Figure 4's function $f(x, y) = y + \sqrt{x + y^2}$ cannot be concavified. This was first proposed by Fenchel (1953) and features in Aumann (1975) and Reny (2010). The function is strictly increasing in both x and y, and strictly concave in each variable individually. It is only weakly quasiconcave, however, because its level sets are straight lines, as shown in the right-hand side of Figure 4. As a result, f is not concavifiable, and, more surprisingly, is not even weakly concavifiable. (continued on the next page)

[2]For reference, the following list of weaker (and easier to prove) relationships between f and g may be useful. (a) If f is strictly concave and g is strictly monotonic, then $g(f(x))$ is not necessarily concave but it is strictly quasiconcave. (b) If f is strictly quasiconcave and g is strictly monotonic, $g(f(x))$ is strictly quasiconcave. (c) If f is strictly quasiconcave and g is weakly monotonic, $g(f(x))$ is at least weakly quasiconcave. (d) If f is weakly but not strictly quasiconcave and g is weakly monotonic, $g(f(x))$ is at least weakly quasiconcave.

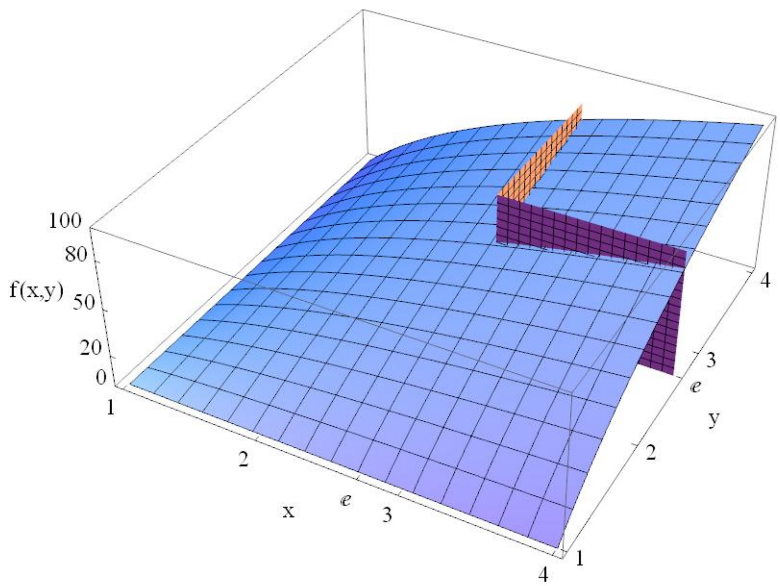

Figure 3: A Strictly Quasiconcave but Not Concave Function $f(x,y) = 50\log(x)\log(y)$ Which Is Concave Only for x and y Greater than e.

(Fenchel's Example, continued) Regardless of any postcomposition by a strictly increasing function g, the rate of increase of the function $g \circ f$ from point a to b is greater than that of the same length segment from d to e. More precisely, the gradient at a is larger than the gradient at c. Hence, if we choose the point b sufficiently close to a, the level sets with values slightly larger than $g(f(a))$ but less than $g(f(b))$, must lie under the segment connecting $(c, g(f(c)))$ to $(b, g(f(b))$ near the point $(c, g(f(c))$. Hence part of the graph of $g \circ f$ lies over the line segment from $(c, g(f(c)))$ to $(b, g(f(b)))$. Hence, $g \circ f$ could not have been concave.

A word of caution regarding this intuitive proof: it depends crucially on the fact that the level sets are only weakly convex. If they were strictly convex, even ever so slightly, then the function f would be concavifiable (see Theorem 4). In that case, g could be chosen in such a way that the gradient would change rapidly so that that the point b must be chosen very close to a in order for the linear gradient approximation to remain valid. However, then the line segment from c to b would lie mostly to the left of the bent level set on which a and c lie, and thus it would not cross, near c, the other level sets in the domain with values greater than $g(f(c))$.

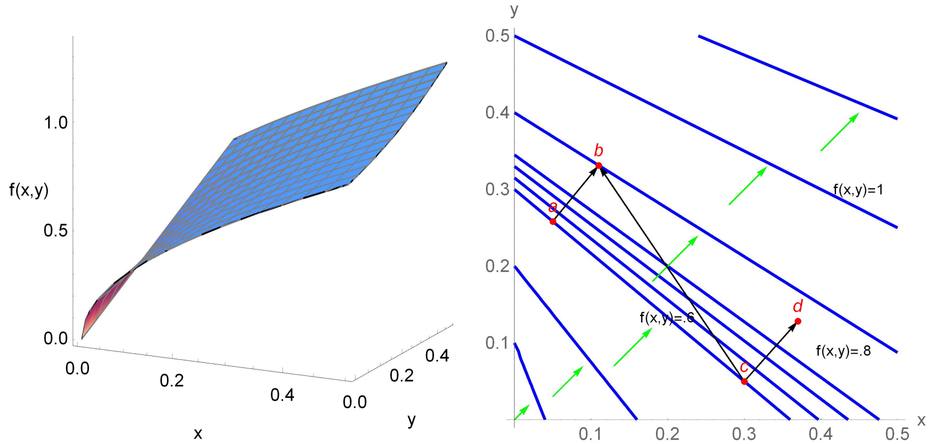

Figure 4: Fenchel's Example: Nonconcavifiable with Linear Level Sets

3 Objective Functions Defined over a Finite Number of Points

Let us start with the easiest case, where the function $f(x)$ is defined on a set consisting of a finite number of points. Afriat's Theorem says that for any finite set of consumption data points satisfying the Generalized Axiom of Revealed Preference (GARP), there exists a continuous, concave, strictly monotone utility function that would generate that data (Afriat [1967]). Another way to view Afriat's Theorem is that if a consumer is solving a maximization problem with a unique solution then he must be maximizing a quasiconcave function, and the theorem says that if we also assume the function is strictly increasing then that function can be chosen to be concave.

The literature following Afriat's Theorem has generalized it to infinite sets of consumption data and in other ways (see Varian (1982), Matzkin & Richter (1991), and Hjertstrand (2011)). Richter and Wong (2004) give three conditions, they call (C),(G), and (G') on a preference ordering over a finite set of consumption bundles. They show that condition (G) is equivalent to being describable by a weakly concave utility function, and (G') is equivalent to being describable by a strictly concave utility function. In fact, condition (G') is equivalent to strict quasiconcavity and condition (C) to weak quasiconcavity, while condition (G) lies between the two. We will present below a significantly shorter proof of their Theorem 2 in the context of functions rather than preference orderings.

Note that in this context the notions of differentiability and continuity are vacuous. Quite different problems arise when the space of points is uncountably infinite (e.g., an interval), as we shall see shortly.

Theorem 1 (Richter and Wong (2004), Theorem 2)**.** *Let $f(\cdot)$ be defined on a finite set of points in $\mathbb{R}^n$. If and only if $f(x)$ is strictly quasiconcave there exists a strictly increasing function $g(z)$ such that $g(f(x))$ is a strictly concave function of x.*

Proof. **Part 1.** We will start by showing that if $f(x)$ is strictly quasiconcave we can find a strictly increasing function $g(z)$ such that $g(f(x))$ is strictly concave.

Define S_1 to be the set of x's that yield the highest value of $f(x)$, S_2 to be the set of that yields the second-highest value, S_3 to yield the third highest, and so forth, so $S_1 \equiv \operatorname{argmax} f(x)$ and $S_i \equiv \underset{x \notin S_1 \cup \cdots \cup S_{i-1}}{\operatorname{argmax}} f(x)$.

In general, the S_i sets may contain many points, though under our assumption that f is strictly quasiconcave, the convex hull W_i of the super-level sets $\cup_{j \leq i} S_i$ will always be a weakly convex polytope in the domain such that $W_i \subset W_{i+1}$ for all i. We really only need to separate these convex polytopes in height in a convex way, but it will be simpler in practice to construct an "overkill" g which convexifies f using pointwise conditions.

To construct our concavified function, we set $g(f(S_1)) = 0$ and choose appropriate numbers ϵ_i inductively such that $g(f(S_{i+1})) \equiv g(f(S_i)) - \epsilon_i$ as follows. Choosing $\epsilon_1 = 1$, we assume $\epsilon_2, \dot{;} \epsilon_{i-1}$ have been chosen by inductive hypothesis. We need to make g have as large a rate of increase as necessary from points in S_{i+1} to S_i to ensure that the rate of increase is always diminishing and g is strictly concave. choose the ϵ_i so that for any $i \geq 2$, and for any $x_{i-1} \in S_{i-1}, x_i \in S_i$ and $x_{i+1} \in S_{i+1}$, we always have,

$$\frac{g(f(x_{i-1})) - g(f(x_i))}{\|x_{i-1} - x_i\|} < \frac{g(f(x_i)) - g(f(x_{i+1}))}{\|x_i - x_{i+1}\|}. \tag{3.1}$$

Since $g(f(x_{i-1})) - g(f(x_i)) = \epsilon_{i-1}$ we can simplify this to choosing ϵ_i so that

$$\epsilon_i > \epsilon_{i-1} \frac{\min_{\substack{x_{i-1} \in S_{i-1} \\ x_i \in S_i}} \|x_{i-1} - x_i\|}{\min_{\substack{x_{i+1} \in S_{i+1} \\ x_i \in S_i}} \|x_i - x_{i+1}\|}. \tag{3.2}$$

From the definition of concavity on arbitrary domains it follows that the choice of ϵ_i in (3.2) guarantees concavity of $g \circ f$.

Part 2. We now must show that if $f(x)$ is not strictly quasiconcave we cannot find a strictly increasing function $g(z)$ such that $g(f(x))$ is strictly concave.

If $f(x)$ is not strictly quasiconcave, then there exist points w, y, z such that $w < y < z$ and one of the following three conditions holds:

(i) $f(y) < Min(f(w), f(z))$ (if $f(x)$ is not even weakly quasiconcave)

(ii) $f(y) = Min(f(w), f(z))$ and $f(w) \neq f(z)$

(iii) $f(y) = f(w) = f(z)$.

In any of these cases, for any strictly monotonic function g, $g(f(y))$ will be below a straight line connecting $g(w)$ and $g(z)$ and hence by Definition 2 is not concave, because:

$$g(f(y)) \leq \left|\frac{w-y}{w-z}\right| g(f(w)) + \left|\frac{y-z}{w-z}\right| g(f(z)) \tag{3.3}$$

In case (i), this is because $g(f(y))$ is less than either $g(f(w))$ or $g(f(z))$, so the inequality in inequality (3.3) is strict. In case (ii), $g(f(y))$ is equal to one of the other two g's and less than the other, so (3.3) is also strict. In case (iii), $g(f(y))$ is equal to both the other two g's, so the inequality becomes an equality. This completes part 2 of the proof. □

Remark 1. Theorem 1 applies equally well to functions defined on finite subsets of arbitrary geodesic metric spaces, if one first generalizes our earlier definitions of quasiconcavity and concavity to the natural analogues of the definitions expanded to entire geodesic metric spaces later in this article. The same proof would apply with the obvious modifications.

Remark 2. If f is only weakly quasiconcave, and so is flat for some values of x, then if the maximum is unique (that is, if argmax $f(x)$ is unique) we can find w, y, z so that case (ii) applies and inequality (3.3) is strict, so that g is not even weakly concave. We will return to this idea of weak quasiconcavity not being definable using weak concavity towards the end of the paper.

Remark 3. If we replace a finite number of points by a countable set of points then Theorem 1 fails. To see this, choose a dense countable set of points, such as the rationals, and apply this to one of our strictly quasiconcave counterexamples (e.g. Example **??** or **??**). The transforming function g would necessarily extend continuously to all of $\mathbb{R}$, providing a contradiction. On the other hand, the proof above also works for domains consisting of an infinite countable set of points which is discrete (so each point is at least a fixed distance ϵ from each other point, e.g. the integers) save perhaps for a unique limit point occurring at the unique argmax of the function f.

Remark 4. Theorem 1 of Richter & Wong (2004) shows when a concave utility function can represent a preference ordering over a finite set of goods (their condition G or E). Kannai (2005) provides an alternative approach to the same problem. Our problem is different because we start with a strictly quasiconcave function rather than with preferences. Thus, we give a simpler answer to a simpler problem, using a particular monotonic postcomposition of the range to concavify a given strictly quasiconcave function.

4 Continuous Objective Functions on $\mathbb{R}^1$

Now let the function's domain be a real continuum. Consider a function $f\colon I \to \mathbb{R}^1$ defined on an interval $I \subset \mathbb{R}^1$ which is either open (in which case it might be unbounded), half open, or closed. For the rest of the paper we will define $\boxed{a \equiv \inf(I)}$ and $\boxed{b \equiv \sup(I)}$. Recall that $\boxed{f^* \equiv \sup(f).}$

The following lemma will allow us to essentially restrict our attention to the case when the postcomposing function g is strictly increasing.

Lemma 1. *Given continuous functions $f\colon D \to \mathbb{R}$ with $D \subset \mathbb{R}^1$ connected and g: $\mathrm{Range(f)} \to \mathbb{R}$, for $g \circ f$ to be strictly quasiconcave it is necessary that f and g fall into precisely one of these four cases:*

(i) f *is strictly increasing and* g *is strictly quasiconcave,*

(ii) $-f$ *is strictly increasing and* $-g$ *is strictly quasiconcave,*

(iii) f *is strictly quasiconcave but not monotone and* g *is strictly increasing, or*

(iv) $-f$ *is strictly quasiconcave but not monotone and* $-g$ *is strictly increasing.*

Proof. The cardinality $\mathrm{card}((g \circ f)^{-1}(t))$ of the level set of value t for $g \circ f$ is $\sum_{s \in g^{-1}(t)} \mathrm{card}(f^{-1}(s))$. In particular, postcomposition by a function never reduces a level set's cardinality. The maximum cardinality of a level set of the quasiconcave function $g \circ f$ is two. Hence on values where the level sets of f have cardinality two, g must be one-to-one, and g can only be two-to-one on values where f has cardinality one. Such continuous functions are fairly simple to analyze.

If all the level sets of f have cardinality one everywhere then f is monotone by continuity and we are in case (i) or (ii). Then g or $-g$ must be quasiconcave since the composition of the strictly increasing function f^{-1} (or $(-f)^{-1}$) with the quasiconcave function $g \circ f$ is again quasiconcave.

If the cardinality of f's level sets is two on a subset $L \subset \mathrm{Range}(f)$, then consider any $p \in L$. Let $\{x, y\} = f^{-1}(p)$, with $x < y$, and let $\{U_i\}$ be any sequence of connected open intervals centered at p and strictly decreasing to p. By continuity of f, the preimages of each U_i under f are open and contain both x and y. Since each point has at most two preimages, for some sufficiently large i, the open subintervals A containing x and B containing y in $f^{-1}(U_i)$ must be distinct, and hence disjoint. If $C = f(A) \cap f(B)$, then f must be monotone on $f^{-1}(C) \cap A$ and $f^{-1}(C) \cap B$. Now C, being connected, either has interior, or else is just p. If it is just p, some points are to the right of x and have values less than p and some points are to the left of y and have values greater than p. Hence, the intermediate value theorem guarantees, since D is connected, another point $z \in (x, y) \subset D$ for which $f(z) = f(x) = f(y) = p$, which violates the cardinality restriction on the level sets. Therefore, we conclude that each point $p \in L$ contains an interval about it, a priori not necessarily in L, on which f is monotone. Since f is one-to-one off of L, f is locally monotone everywhere except for local extrema. The cardinality rule then implies that f has at most two extrema and is monotone on each connected segment after removing these points.

Now g preserves the local extreme points of f. So if f has more than one local extreme point, then the values must coincide under g, but then there are at least four preimages of some value near this extreme value for $g \circ f$, violating strict quasiconvexity. So f can have at most one local extremum. It has exactly one because f is not monotone and D is connected. In particular, L is connected.

Since L is connected, g is one-to-one on the closure of L, not just on L. This contains all of the local extrema of f. Now assume g is strictly increasing on this closure. If the graph of g changes direction elsewhere, then $g \circ f$ has at least two local extrema, violating quasiconcavity. Hence, g is strictly increasing everywhere. This then implies that f was strictly quasiconcave since postcomposition by a strictly increasing function g^{-1} preserves quasiconcavity.

If g was strictly decreasing on the closure of L then similarly it is strictly decreasing everywhere and $-f$ is strictly quasiconcave. □

Remark 5. Lemma 1 fails for discontinuous f and g or for disconnected D. For instance, f can be pathologically discontinuous and not monotone, and yet one-to-one so $g = f^{-1}$ exists and is one-to-one with the identity function as $g \circ f$.

Remark 6. Surprisingly, the analogous lemma for weakly quasiconcave compositions is false. If we replace "strictly" everywhere by "weakly", however, and "strictly increasing" by "nondecreasing", then the conclusion remains true provided we conclude that f or $-f$ is weakly quasiconcave only after collapsing regions in the range where g is constant.

4.1 The Case Where the Objective Function on $\mathbb{R}^1$ Is Continuous and Strictly Monotone

If the continuous function f is strictly increasing or decreasing, then $f\colon I \to \mathbb{R}^1$ is invertible. Hence, we can easily solve the problem of concavifying f by choosing $g = h \circ f^{-1}$ where h is a concave function and hence $g(f(x)) = h(x)$ is concave.[3]

Here, however, we will treat the twice-differentiable case more intrinsically and connect the definition of concavity more viscerally with the properties of f. This will build a foundation for the next section, where we treat the noninvertible case.

Thus, let us move to the special case where f is twice differentiable and strictly increasing. We will also for now assume g is twice differentiable and strictly increasing. Recall that we are searching for a strictly concave function $g \circ f$. The twice differentiable function $g \circ f$ is strictly concave if the expression

$$(g \circ f)''(x) = g''(f(x)) \cdot f'(x)^2 + g'(f(x)) \cdot f''(x), \tag{4.1}$$

is nonpositive and never vanishes on an entire interval. Equivalently, this occurs if

$$\frac{g''(f(x))}{g'(f(x))} \leq -\frac{f''(x)}{f'(x)^2}, \tag{4.2}$$

with equality never holding on any interval. (There are examples of strictly convex functions where equality holds on a full measure set, however.)

Since we assume that f is strictly increasing, it is invertible. Define $\boxed{z \equiv f(x)}$, so $x = f^{-1}(z)$. Note that $\frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(z))} = \frac{\partial}{\partial z} f^{-1}(z)$, so the right-hand side of inequality (4.2) is

$$-\frac{f''(x)}{f'(x)^2} = \frac{\partial}{\partial x}\left(\frac{1}{f'(x)}\right) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial(z = f(x))} f^{-1}(f(x))\right) = f^{-1\prime\prime}(f(x)) \cdot f'(x). \tag{4.3}$$

Taking another step:

$$-\frac{f''(x)}{f'(x)^2} = f^{-1\prime\prime}(f(x)) \cdot f'(x) = \frac{\partial}{\partial x} \log f^{-1\prime}(f(x)). \tag{4.4}$$

Similarly (though since we are constructing g we need not invert), the left hand side of inequality (4.2) is $\frac{\partial}{\partial f(x)} \log g'(f(x))$. Hence a sufficient criterion

[3] This can even be done in the case that f is discontinuous and strictly increasing or decreasing, by first repairing the range (see Section 6.1). However, note that basic obstructions like those in Example 2 arise in two or more dimensions regardless of continuity.

for inequality (4.1) to be true, so that we do not have to even worry about it vanishing on an interval, is that

$$\frac{\partial}{\partial z}\log g'(z) < \frac{\partial}{\partial z}\log f^{-1\prime}(z) \tag{4.5}$$

for all z in the range of f, provided both sides are well defined.

If we choose a number $c > 0$ and a function g so that

$$g'(z) = e^{-cz} \cdot f^{-1\prime}(z) \tag{4.6}$$

then

$$\log g'(z) = -cz + \log f^{-1\prime}(z) \tag{4.7}$$

and

$$\frac{\partial}{\partial z}\log g'(z) = -c + \frac{\partial}{\partial z}\log f^{-1\prime}(z) < \frac{\partial}{\partial z}\log f^{-1\prime}(z) \tag{4.8}$$

Integrating equation (4.6)'s g' will produce the desired function.

This approach, using Equation (4.6), has the advantage of relying only on first derivative data of f in its construction (although the properties requisite for its application do rely on the second derivatives of f.) Later we will make use of a second approach. To construct g, choose any function $u(z) < f'(f^{-1}(z)) \cdot f^{-1\prime\prime}(z)$. Then set

$$g(z) = \int_0^z \left(e^{\int_a^s u(t)dt} \right) ds \tag{4.9}$$

for any $a \in \mathbb{R}$ chosen at the top of the domain of $u(z)$. This yields a function with $g \circ f$ concave on $(a, d]$.

We note that demanding strict equality in quantities such as (4.5) does not lose us any generality, provided we are willing to be flexible with g. Since if $g \circ f$ is strictly convex, by a modification of g as above we can make its second derivative strictly negative.

4.2 The Case When the Objective Function on $\mathbb{R}^1$ Is Nonmonotonic but Twice Differentiable

Now suppose the function f is twice differentiable and strictly quasiconcave but not monotone. In that case it achieves its maximum at a unique internal point $m \in (a, b)$, so that f is rising on $(a, m]$ and falling on $[m, b)$ as in Figure 5.

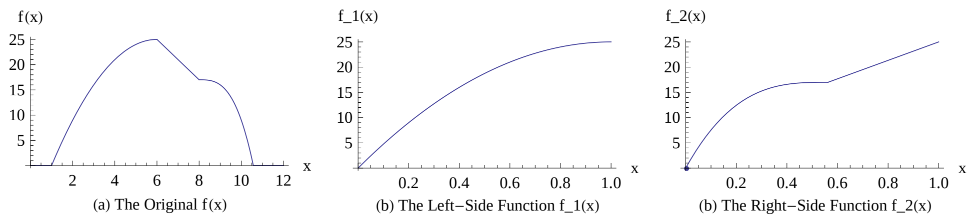

Figure 5: The Construction of f_1 and f_2

For now, we will also require that f have a non-vanishing derivative except at an internal maximum or endpoints.

Denote by $\boxed{f_1:}$ $[0, 1) \rightarrow \mathbb{R}$ the strictly increasing function

$$f_1(x) = f(a(1 - x) + xm)$$

and by $\boxed{f_2:}$ $[0, 1) \rightarrow \mathbb{R}$ the strictly increasing function

$$f_2(x) = f(b(1 - x) + xm).$$

Figure 5 illustrates this construction, which splits $f(x)$ into two strictly increasing functions on $[0, 1]$ to save the bother of using negative signs and absolute values of slopes in our analysis. (In Figure 5 the $f(x)$ is not twice differentiable; it is drawn with a kink to illustrate the f_1, f_2 construction clearly.)

The new functions f_1 and f_2 are homeomorphisms onto their images, so they have inverses f_1^{-1} and f_2^{-1}. Hence, by post-composition we can easily choose a g such that either $g \circ f_1$ or $g \circ f_2$ is strictly concave and smooth. The difficulty is in making $g \circ f$ concave on its entire domain— that is, to use the same function to concavify both f_1 and f_2— especially when f is nondifferentiable or not defined over a compact set. We will treat this general problem in the next section.

Before handling the present case, with $f(x)$ being twice differentiable, we will present an example demonstrating why we need the condition that the gradient not vanish except at the maximum. Note first that since the gradient is continuous, the derivative is bounded away from zero except perhaps at the optimum and endpoints.

Figure 6s example: A zero right-derivative. A nondifferentiable example is the simplest way to see the problem. Consider the strictly quasiconcave function $f(x)$ in Figure 6, which is defined as follows.

$$f(x) = \begin{cases} x & \text{if } x \leq 1 \\ 1 + (x-1)^2 & \text{if } 1 \leq x \leq 2 \\ 4 - x & \text{if } x \geq 2 \end{cases}$$

The problem in concavifying $f(x)$ comes with the x values around 1 and 3. We have $f(1) = f(3) = 1$, so necessarily $g(f(1)) = g(f(3))$. But $f'_+(1) = 0$ and $f'(3) = -1$, which makes it impossible for $g(f(x))$ to be concave since either $g \circ f'(1)$ remains 0 or else $g \circ f'(3)$ becomes infinite.

xxxx[This figure ahs a strange parallel horizontal line at the bottom. It needs bigger fonts too.]

The same kind of problem shows up with differentiable strictly quasiconcave functions.

Figure 7s example: An inflection point. Consider the strictly quasiconcave function $f(x) = -x^4 + 6x^3 - 12x^2 + 10x$ in Figure 7. The problem in concavifying $f(x)$ comes with the x values around 1 and 3. We have $f(1) = f(3) = 3$, so necessarily $g(f(1)) = g(f(3))$. But $f'(1) = 0$ and $f'(3) = -8$, which makes it impossible for $g(f(x))$ to be concave since either $(g \circ f)'(1) = 0$ or else $(g \circ f)'(3) = -\infty$. Thus, the $(g \circ f)$ shown in Figure 7 is not concave.

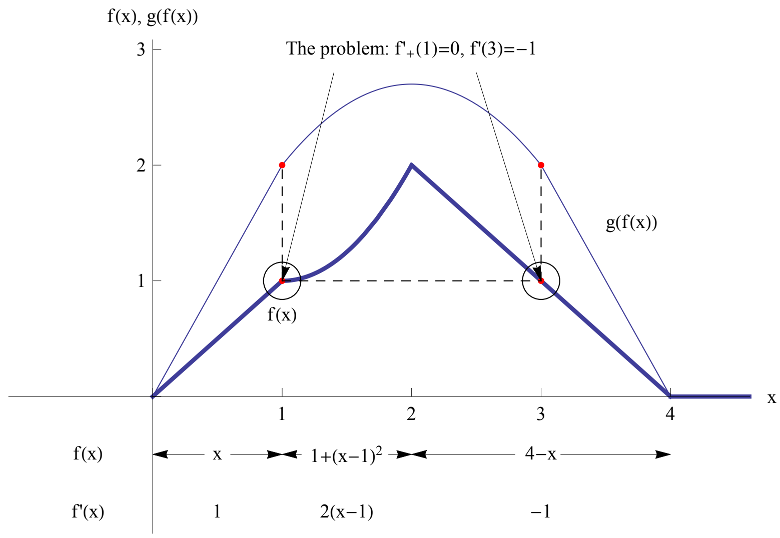

Figure 6: A Nonconcavifiable Strictly Quasiconcave Function with a Zero Right-Derivative at $x = 1$

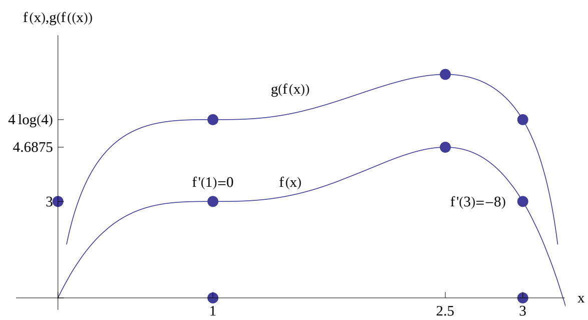

Figure 7: A Nonconcavifiable Strictly Quasiconcave Function with a Zero Derivative at $x = 1$

Since f is twice differentiable with first derivative bounded away from 0, and I is compact, the problem becomes easy in light of what we discovered in the previous section. Simply set

$$\text{(Concavifying function)} \quad g'(z) = e^{\int_0^z u(t)dt} \cdot f_1^{-1\prime}(z) \cdot f_2^{-1\prime}(z) \tag{4.10}$$

for any continuous function u: $\mathbb{R} \to \mathbb{R}$ satisfying

$$u(z) < \min\left\{0, -\frac{\partial}{\partial z}\log f_1^{-1\prime}(z), -\frac{\partial}{\partial z}\log f_2^{-1\prime}(z)\right\}. \tag{4.11}$$

so that

$$\frac{\partial}{\partial z}\log g'(z) = \frac{\partial}{\partial z}\log {f_1^{-1}}'(z) + \frac{\partial}{\partial z}\log {f_2^{-1}}'(z) + u(z) < \min\left\{\frac{\partial}{\partial z}\log {f_1^{-1}}'(z), \frac{\partial}{\partial z}\log {f_2^{-1}}'(z)\right\}. \tag{4.12}$$

Note that we used the nonvanishing derivative condition on f simply to guarantee the existence of u in that the right-hand side of (4.11) is bounded from below. Thus we can solve for g, yielding $g \circ f$ concave.

We have saved what we consider to be the most beautiful example of this paper for last. It demonstrates the most subtle type of obstruction to concavifiability that can arise.

Figure 8s example: A positive log-derivative with unbounded variation. Consider the strictly quasiconcave function $f(x)$ on $[-1,4]$ shown in Figure 8, which is defined as follows.

$$f(x) = \begin{cases} q(x) & -1 \le x < 1 \\ q(1) - \frac{1}{2}(x-3)(x-1)q'(1) & 1 \le x \le 4 \end{cases}$$

where

$$q(x) = \int_{-1}^{x} e^{t\sin\left(\frac{1}{t}\right)+1}\, dt$$

From the formula we can readily verify that the first derivative,

$$f'(x) = \begin{cases} e^{x\sin(1/x)+1} & -1 \le x < 1 \text{ and } x \neq 0 \\ e & x = 0 \\ e^{1+\sin(1)}(2-x) & 1 \le x \le 4 \end{cases},$$

is a C^1 function with derivative bounded away from 0, except at the peak of f at $x = 2$, and is strictly concave on $[1,4]$. Nevertheless, on the interval $(-1,1)$, we have $\log f'(x) = x\sin(\frac{1}{x}) + 1$ which is a classic example of a function with unbounded variation. It is not obvious at this point, but later in the paper, Theorem 2 will prove that such an f cannot be concavified by any postcomposition.

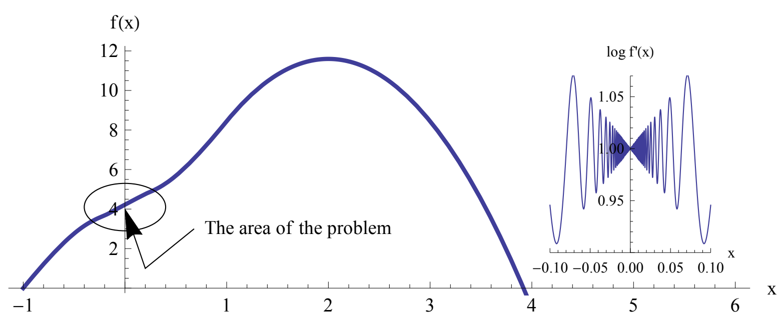

Figure 8: Most Subtle of All— Nonconcavifiable Strictly Quasiconcave Function with Strictly Positive Derivatives but Unbounded Variation

The problem in Figure 8s example is that the log of the first derivative has infinite variation over a finite interval, is imperceptible to the naked eye when examining the graph. In trying to concavify this, it becomes impossible to adjust the slope to be decreasing at a certain height on one side of the maximum without creating a vertical or horizontal tangency at the same height on the other side. We will show this formally later, as part of Theorem 2. Note that if f is monotonic— as it would be if f were restricted to the range $[-1, 2]$— then unbounded variation does not hinder concavifiability

4.3 The Case When the Function on $\mathbb{R}^1$ Is Nondifferentiable and Nonmonotonic but Continuous

What is most difficult is when f is nondifferentiable and nonmonotonic. If after we postcompose f_2 with f_1^{-1}, the new function becomes smooth except at the endpoints of the domain, then we are simply in the case handled by the previous section. This will generally not hold true, however.

We will begin by showing that the most obvious avenues of approach to constructing a concavifying function g fail.

If we could arrange for g to smooth f before it concavifies f, f's nondifferentiability wouldn't matter. Example X shows that this approach won't work.

Example X: You can't always smooth.
One cannot always smooth a non-differentiable quasiconcave f by postcomposition with a $g\colon \mathbb{R} \to \mathbb{R}$, except by one which is constant on the range of f. First, choose a smooth function $q_1\colon [0, 1) \to \mathbb{R}$ with strictly positive derivative (e.g. $q(x) = x$) and a continuous strictly increasing function $q_2 :\ [0, 1) \to \mathbb{R}$ with the same range as q_1 but nondifferentiable at a countable dense set of points in $[0, 1]$. [4] Now, form the strictly quasiconcave $f\colon (-1, 1) \to \mathbb{R}$ with $m = 0$, $f_1 = q_1$ and $f_2 = q_2$. For any $g\colon \mathbb{R} \to \mathbb{R}$, if $g \circ f$ restricted to $(-1, 0]$ is smooth then g is smooth on the entire range of f. Consequently $g \circ f$ is not differentiable on a dense subset of $(0, 1)$.

[4] One can build such a q_2 by repeatedly modifying a Cantor staircase function, for example. (See http://rasmusenorg/papers/dense.pdf.

[dense.pdf is not up on the web—put it there or remove the link.]

A second consideration is that in our search for a suitable g we cannot expect g to always be concave, even to concavify a single strictly increasing function, or g' to be even locally Lipschitz— that is, to be such that on each compact subset of the domain D, $\sup_{\substack{x,y\in D \\ y\neq x}} \frac{|g(y)-g(x)|}{|y-x|} < C$ for some $C > 0$. Example Y below shows this using the fact that concave functions are differentiable except at a countable number of points (in fact, for strictly increasing concave functions the left- and right-hand derivatives exist everywhere and are both nonincreasing.) In particular, concave functions are locally Lipschitz, so their slopes are bounded.

Example Y: Concavifiers of Lipschitz f need not be Lipschitz. Suppose f: $[a,b] \rightarrow \mathbb{R}$ is a strictly increasing function which is differentiable except possibly at $c \in (a,b)$ and which has a horizontal tangent at c, i.e. $\lim_{t\mapsto c^-} f'(x) = 0$. A Lipschitz example is $f(x) = x^3$ confined to [-1,1], for which $f'(0) = 0$. This implies that f^{-1} has the opposite corner with $\lim_{t\mapsto a^-} f^{-1\prime}(x) = \infty$, and in particular that f^{-1} is not Lipschitz. Any concavifying g can be written as f^{-1} followed by a locally Lipschitz concave function. Since f^{-1} is not Lipschitz any g concavifying f cannot be locally Lipschitz, and thus cannot be concave, unless g is constant on an interval containing $[f(c), f(b)]$. Similarly, if there is an interior point $w \in (a,b)$ for which there is a vertical tangent, $\lim_{t\mapsto w^+} f'(x) = \infty$, then f^{-1} has a horizontal tangent at $f(w)$. Hence, any concavifying g cannot itself be concave unless it is constant on $[f(a), f(w)]$.

In the case of a nondifferentiable f, we would like, following equation (4.10), to form

$$g'(z) = e^{\int_0^z u(t)dt} f_1^{-1\prime}(z) \cdot f_2^{-1\prime}(z) \tag{4.13}$$

in a distributional sense, since we can only rely on weak derivatives.[5] For this we consider the Sobolev space $W^{k,p}$, the space of functions whose weak k-th derivatives belong to L^p. Since f_1^{-1} and f_2^{-1} are strictly increasing, they are absolutely continuous and live in $W^{1,1}$. However, $W^{1,1}$ does not form

[5]A weak derivative is a generalization of the concept of the derivative to nondifferentiable functions. See expression (5.10) and surrounding text.

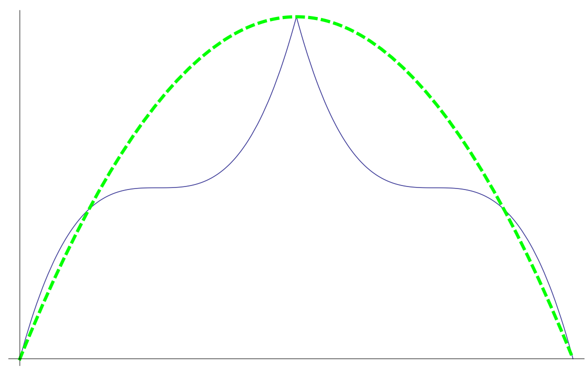

Figure 9: A Function Made from Back-to-Back Cubic Functions

an algebra, since it is not closed under multiplication of functions, and this creates the problem shown in Figure 9s example.

> **Figure 9s example: No simple concavifier.** Suppose our quasiconcave function $f(x)$ was such that $f_1(x) = f_2(x) = x^3$, as in Figure 9. The strictly increasing function $f_1^{-1}(x) = f_2^{-1}(x) = x^{\frac{1}{3}}$ belong to $W^{1,1}$ on $[-1, 1]$. This has derivative $f_1^{-1\prime}(x) = \frac{1}{3}x^{-\frac{2}{3}} \in L^1$, but the product we would have for our construction in equation (4.13) is $f_1^{-1\prime}(z) \cdot f_2^{-1\prime}(z) = \frac{1}{9}x^{-\frac{4}{3}}$, which is not in L^1, and integrating it to get g yields $-\frac{1}{3}x^{-\frac{1}{3}}$, which is not even increasing on all of $[-1, 1]$. On the other hand, this f is easily concavified by $g(y) = -y^{2/3}$.

Figure 9s example shows that simply taking the product $f_1^{-1\prime}(z) \cdot f_2^{-1\prime}(z)$ for $g'(z)$ will not always work. If we assumed that each derivative was in $W^{1,p}$ for $p \geq 2$, then the product would be in $W^{1,1}$. If we want to use arbitrary products, though, we would be forced to work in $W^{1,\infty}$, and this is a stronger assumption than we need, since $W^{1,\infty}$ coincides with the space of Lipschitz functions and we know that there are non-Lipschitz quasiconcave functions that can be concavified.[6]

Conversely, if we wanted to work directly on weak second derivatives to guarantee that $\frac{\partial}{\partial z} \log g'(z) < \min\left\{\frac{\partial}{\partial z} \log f_1^{-1\prime}(z), \frac{\partial}{\partial z} \log f_2^{-1\prime}(z)\right\}$ for equation (4.12), we would need to work in $W^{2,1}$. By the Sobolev embedding theorem, however, $W^{2,1} \subset W^{1,\infty}$ for one-dimensional functions.[7] Thus we gain nothing

[6]Consider $f(x) = x^{\frac{1}{3}}$, which has an unbounded first derivative, $f'(x) = \frac{1}{3}x^{-\frac{2}{3}}$.

[7]For the general Sobolev Embedding Theorem, see Wikipedia (2011) or Chapter 2 of

over working with Lipschitz functions.

In view of these problems, let us consider the following upper and lower derivatives $\overline{D}f, \underline{D}f \colon \mathbb{R} \to [-\infty, \infty]$ (not to be confused with "weak derivatives"). Given a function f define

$$\overline{D}f(x) = \limsup_{h \to 0} \frac{f(x+h) - f(x)}{h} \quad \text{and} \quad \underline{D}f(x) = \liminf_{h \to 0} \frac{f(x+h) - f(x)}{h}. \tag{4.14}$$

These quantities always exist, if we allow for values of $-\infty$ and ∞, and $\overline{D}f(x) \geq \underline{D}f(x)$ with equality occurring if and only if the derivative of f exists at x, in which case both quantities coincide with $f'(x)$.

Note that if $w < x < y$ then the slope of the secant line between $(w, f(w))$ and $(y, f(y))$ lies between the values of the slopes of the secant lines from $(w, f(w)))$ to $(x, f(x))$ and $(x, f(x))$ to $(y, f(y)))$. Hence we have,

$$\overline{D}f(x) = \limsup_{\substack{|y-w| \to 0 \\ w \leq x < y}} \frac{f(y) - f(w)}{y - w} \quad \text{and} \quad \underline{D}f(x) = \liminf_{\substack{|y-w| \to 0 \\ w \leq x < y}} \frac{f(y) - f(w)}{y - w}. \tag{4.15}$$

In other words, these quantities reflect the lower and upper limit of the slopes of all secant lines between points before and after x, not just those with an endpoint at x. Also, we cannot dispense with the ordering $w \leq x < y$ in the above limits, since the continuous extension of the function $x^2 \sin(\frac{1}{x^2})$ has derivative 0 at 0 and yet admits secant lines of unboundedly positive and negative slope whose endpoints are arbitrarily close to 0.

We now begin to explore analogues of conditions for concavity of C^2 functions using the above objects that are available to us for arbitrary continuous functions. For what follows let $\ell(x, y)$ represent the secant line between $(x, f(x))$ and $(y, f(y))$, and let $s(x, y)$ represent the slope of $\ell(x, y)$. In the next three lemmas, we will explore the relationship between concavity and conditions on $\overline{D}f$. (Analogous statements invoking $\underline{D}(f)$ could also naturally be formulated.)

Lemma 2. *A continuous function $f : I \to \mathbb{R}$ is strictly concave if and only if for all $x \in I$, and all $w < x$, $\overline{D}f(x) < s(w, x)$.*

Proof. If f is concave, then for any $y > w$ in I we have $s(x, y) < s(x, w)$ and hence taking the lim sup as $y \to x$ we obtain the forward implication.

Aubin (1982). For $W^{2,1}(\mathbb{R}) \subset W^{1,\infty}(\mathbb{R})$, see Exercise 2.18 in Ziemer (1989).

Conversely, if f is not convex then there exist points $r < s < t$ in I such that $f(s)$ lies below the secant line $\ell(r,t)$. By continuity, one may trace the graph in both directions from $(s, f(s))$ until it runs into the segment $\ell(r,t)$ showing that there is some open interval $(w,x) \subset (s,t)$ such that the entire graph of f over (w,x) lies strictly below the secant line $\ell(w,x)$. For all $z \in (w,x)$, we then have $s(z,x) > s(w,x)$ and therefore $\overline{D}f(x) \geq s(w,x)$ contradicting our hypothesis. $\square$

Lemma 3. *A continuous function $f : I \to \mathbb{R}$ is strictly concave if and only if $\overline{D}f$ is a strictly decreasing function.*

Proof. If f is concave, then for any three points $w < x < y$ in I we have $s(w,x) > s(w,y)$ and $s(w,y) > \overline{D}f(y)$ by Lemma 2. Taking the lim sup as w approaches x from below we see that $\overline{D}f(x) \geq s(w,y) > \overline{D}f(y)$, as desired.

Conversely, if f is not convex we can find, as in Lemma 2s proof, points $w < x$ for which the entire graph of f over (w,x) lies strictly below the secant line $\ell(w,x)$. After possibly shrinking this neighborhood we may assume the graph of f changes sides of the secant line $\ell(x,w)$ at both w and x. Then for any point z sufficiently near w, and any point y sufficiently near x, we have $s(z,w) < s(w,x) < s(x,y)$. Taking limsup's as z approaches w and y approaches x, we obtain $\overline{D}f(w) \leq \overline{D}f(x)$, contradicting our hypothesis. $\square$

Combining Lemma 3 with Figure 6s example, we obtain the following necessary criterion for strict quasiconcavity of f. First recall that a was defined as the lower bound of fs support and m as its argmax in Section 4.2. We will use the following notation in the next proof and the remainder of the paper: $\boxed{f\rfloor_{(a,m]}}$ denotes the restriction of the function f to the interval $(a,m]$.

Lemma 4. *Given a strictly quasiconcave function $f\colon (a,b) \to \mathbb{R}$, there is a $g\colon \mathbb{R} \to \mathbb{R}$ such that $g \circ f$ is strictly concave only if there is a function $h\colon \mathbb{R} \to \mathbb{R}$ such that $h \circ f$ satisfies,*

$$\begin{cases} 0 < \underline{D}(h \circ f)(x) \leq \overline{D}(h \circ f)(x) < \infty & \text{for } x \in (a,m) \\ \\ -\infty < \underline{D}(h \circ f)(x) \leq \overline{D}(h \circ f)(x) < 0 & \text{for } x \in (m,b). \end{cases} \tag{4.16}$$

In particular, $h \circ f$ must be (locally) Lipschitz except at m.

Proof. Suppose there is no such h. We can compose f by the function $h = f^{-1}_{\downharpoonright_{(a,m]}}$ so that $h \circ f$ is still strictly quasiconcave, but $(h \circ f)_1$ is linear. By hypothesis, $(h \circ f)_2$ either admits a vertical tangency on the pre-image of the range of $(h \circ f)_1$, or else $\underline{D}(h \circ f)_2(x) = 0$ for some $x \in (0,1)$. In the latter case, if we had instead chosen $h = -f^{-1}_{\downharpoonright_{[m,b)}}$ then $(h \circ f)_2$ would be linear and $(h \circ f)_1$ would admit a vertical tangency, and so we are back in the first case after switching "1"and "2". Hence, without loss of generality we may assume that there is a point $x \in (0,1)$ with $\overline{D}(h \circ f)_2(x) = \infty$. (Recall here that $(h \circ f)_2$ is increasing, see Figure 5.)

Since $(h \circ f)_1$ is the identity, any strictly concavifying g for $h \circ f$ must be concave and strictly increasing, and hence with $\underline{D}(g)(z) > 0$ for any z in the interior of the range of $(h \circ f)_1$. Then, however, it could not have concavified $(h \circ f)_2$. $\square$

The function h in Lemma 4 can also be taken to be the inverse of the restriction to the strictly increasing side, so $h = f^{-1}_{\downharpoonright_{[m,b)}}$.

Remark 7. In fact, we shall see that the necessary conditions in Lemma 4 turn out to not be sufficient for concavifiability. A significantly more subtle problem arises.

We will need terminology to be able to discuss the problem of rapidly changing derivatives.

$\boxed{\text{The "variation" of a function}}$ $f \colon [a,b] \to \mathbb{R}$ is defined as

$$\boxed{\mathrm{Var}(f)} \equiv \sup \left\{ \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})| \colon n \in \mathbb{N} \text{ and } a \leq x_0 < x_1 < x_2 < \cdots < x_n \leq b \right\}. \tag{4.17}$$

Denote the "functions of bounded variation" on the closed interval $[a,b]$ by

$$\boxed{\mathrm{BV}([a,b])} \equiv \{ f \colon [a,b] \to \mathbb{R} \colon \mathrm{Var}(f) < \infty \}. \tag{4.18}$$

For a general interval $I \subset \mathbb{R}$, denote by $\boxed{\mathrm{BV}_{loc}(I)}$ the set of *locally* bounded variation functions, i.e. those which belong to $\mathrm{BV}([a,b])$ for every compact interval $[a,b] \subset I$.

In Theorem 2 below, without loss of generality, we will assume that our function $f \colon (a,b) \to \mathbb{R}$ has the property that if either f or $-f$ is quasiconcave, then $\text{Range}(f_{\rfloor_{(a,m]}}) = \text{Range}(f)$ where m represents the unique extremal point. [8]

Thus armed with notation, we can state our second theorem. Recall that we earlier explained that for monotonic functions it is very easy to show that the function being strictly quasiconcave is equivalent to it being concavifiable. For nonmonotonic functions we need to add two more conditions whose importance is considerably less clear.

Theorem 2. *For any continuous nonmonotonic function $f : (a,b) \to \mathbb{R}$, there is a function g: $\text{Range(f)} \to \mathbb{R}$ such that $g \circ f$ is strictly concave if and only if*

(i) f or $-f$ is strictly quasiconcave, with argmax m.

(ii) for $h \equiv f^{-1}_{\rfloor_{(a,m]}}$, the function $h \circ f_{\rfloor_{(m,b)}}$ and its inverse are locally Lipschitz.

(iii) $\log\left|\overline{D}(h \circ f)\right| \in \text{BV}_{loc}((m,b))$.

Moreover, when g exists it is strictly monotone.

Proof. We will first explain the necessity of (i), strict quasiconcavity. Apply exactly the same proof as Part 2 of Theorem 1s proof (the theorem for functions defined over a finite number of points). Here, f is defined over a continuum, but we can still apply the proof method of choosing three points from the continuum to test for concavity. Lemma 4 demonstrates the necessity of condition (ii) on the upper derivatives. Condition (iii) will be explained at the end of the proof.

The sufficiency of the conditions is a bit more difficult to prove. Without loss of generality we can assume f is strictly quasiconcave. The existence of g implies that it is continuous, since f and $g \circ f$ are. Moreover, $\overline{D}(h \circ f)$ is nonpositive on (m,b) since f is strictly decreasing there and h is strictly increasing on its domain. (This accounts for taking the absolute value in condition (iii), which is unnecessary in the case when $-f$ is quasiconcave.) Since $\log\left|\overline{D}(h \circ f)\right| \in \text{BV}_{loc}((m,b))$, it is a standard fact, e.g. see Folland (1984), that $\log\left|\overline{D}(h \circ f)\right|$ is the difference of two strictly in-

[8]If this is not the case, we simply replace $f(x)$ by its reflection about $\frac{b+a}{2}$, namely $f(b+a-x)$. Then the very same resulting function g will concavify the original f.

creasing functions. Also, $\log\left|\overline{D}(h\circ f)\right|$ is continuous except at a countable set of points such that the sizes of the jumps at the discontinuities on any compact interval are summable. Now $\log|(h\circ f)'|$ agrees with this function wherever it is defined, which is a-priori almost everywhere. [xxx Eric is confused by this last clause]By the DArboux [xxx Chris check spelling] Theorem (Folland (1984)), and since a full measure set is dense, $\log|(h\circ f)'|$ then agrees with $\log\left|\overline{D}(h\circ f)\right|$ at each point where $\log\left|\overline{D}(h\circ f)\right|$ is continuous. Since $-\log|(h\circ f)'(x)| = \log|((h\circ f)^{-1})'(h\circ f(x))|$, it also agrees with $\log\left|\overline{D}(h\circ f)^{-1}\circ h\circ f\right|$ except at a countable number of points where $\log\left|\overline{D}(h\circ f)^{-1}\circ h\circ f\right|$ has discontinuities with summable gaps. Thus, since precomposition does not affect the BV property, except for the domain over which it applies, $\log\left|\overline{D}(h\circ f)^{-1}\right| \in \mathrm{BV}_{loc}\left((h(f(m)), h(f(b)))\right)$.

Let $h_1 \equiv h_0\circ h$, where h_0 is a smooth strictly increasing concave function on $(-\infty, f(m)]$ with $\lim_{x\to m}\overline{D}(h_0\circ f)(x) = 0$. This can always be done by using an h_0 that increases sufficiently slowly near $f(m)$.

From now on in the proof, we will write f for $f_{\rfloor[m,b)}$ to avoid distraction from the subscript. Since the derivative of h_0 is bounded away from 0 and ∞ and is strictly decreasing on any compact subinterval of $(f(b), f(m))$, the function $\log\left|\overline{D}((h_1\circ f)^{-1})\right|$ still lies in $\mathrm{BV}_{loc}((h_1(f(b)), h_1(f(m))))$ and $h_1\circ f$ is concave on $(a, m]$.

Now choose $z_0 \in (h_1(f(b)), h_1(f(m)))$. Since $\log\left|\overline{D}((h_1\circ f)^{-1})\right| \in \mathrm{BV}_{loc}\left((h_1(f(b)), h_1(f(m)))\right)$, there is a representative

$$q \in L^1_{loc}\left((h_1(f(b)), h_1(f(m)))\right) \tag{4.19}$$

of the almost everywhere defined function $\left(\log\left|\overline{D}((h_1\circ f)^{-1})\right|\right)'$ such that $\log\left|\overline{D}((h_1\circ f)^{-1})(z)\right| = \log\left|\overline{D}((h_1\circ f)^{-1})(z_0)\right| + \int_{z_0}^{z} q(t)\,dt$. Since $\lim_{z\to m}\overline{D}(h_0\circ f)(z) = 0$, the negative part of q, defined by

$$q^-(x) \equiv \begin{cases} q(x) & q(x) < 0 \\ 0 & q(x) \geq 0 \end{cases}, \tag{4.20}$$

belongs to $L^1([h_1(f(c)), h_1(f(m))])$ for any $c < b$. By integrating, we can find a twice- differentiable (though not necessarily in C^2) function g: $(h_1(f(b)), h_1(f(m))) \to \mathbb{R}$ such that

$$g'(z) = e^{\int_{h_1(f(m))}^{z}(-1+q^-(t))\,dt}. \tag{4.21}$$

Consequently, $g'(z) > 0$, $g''(z) < 0$ and $(\log g')' < q^-(z)$ for each $z \in (h_1(f(b)), h_1(f(m)))$. Since by construction $(\log g')'(z) < (\log|((h_1\circ f)^{-1})'|)'(z)$

for almost every $z \in (h_1(f(b)), h_1(f(m)))$, it follows that $\overline{D}(g \circ h_1 \circ f)$ is strictly decreasing on (a, b) and so $g \circ h_1 \circ f$ is concave, which is what we needed to prove.

All that remains to be proved is the necessity of condition (iii). If $\log \overline{D}(h \circ f) \notin \mathrm{BV}_{loc}((m, b))$, then no such function $q \in L^1$ can be found: there exists no g for which $\log g'$ grows slower than $\log(h_1 \circ f)'$ since $\log g'(z)$ would necessarily become unbounded before z reached $h_1(f(b))$. $\square$

Theorem 2's condition (ii) is, strictly speaking, superfluous in that it only serves to establish the existence of the function in condition (iii), where its existence is implicit. In particular, we need $\log \left|\overline{D}(h \circ f)\right|$ to exist almost everywhere in order to make sense of it being in BV_{loc}. Once it belongs to BV_{loc} we can conclude that $h \circ f$ on (m, b), and its inverse, are locally Lipschitz. Moreover, whenever the hypotheses of the Theorem 2 hold, g will always be strictly monotone.

This theorem shows that quasiconcavity is not quite equivalent to concavifiability. Besides quasiconcavity we need condition (ii), which roughly says that after straightening out one side, the other side has no horizontal or vertical tangencies. Beyond that, one still needs the yet more subtle condition (iii) governing the oscillation of the derivative on the unstraightened side. In economic terms, the marginal utility or profit cannot be oscillating too wildly. Note that even after straightening one side to linear, f can have a horizontal or vertical tangency at m or b on the unstraightened side and still be concavifiable, as in Figure 10. Also, any pathology is permitted on one side so long as it is mirrored on the other side, as with Figure 9's symmetric zero slopes at inflection points.

Remark 8. The statement for functions of the form $f\colon (a, b] \to \mathbb{R}$ (that is, where the domain is $(a, b]$, not (a, b)), is identical, except that $h \circ f\rfloor_{(m,b]}$ should be locally Lipschitz and $\log \overline{D}(h \circ f) \in \mathrm{BV}_{loc}((m, b])$. For the remaining two cases of possible interval domains, $f\colon [a, b) \to \mathbb{R}$ and $f\colon [a, b] \to \mathbb{R}$, the statement is identical to the original or the modification except that h becomes $h = \left(f\rfloor_{[a,m]}\right)^{-1}$. The necessary modifications of Theorem 2's proof for these three cases are straightforward. The only nontrivial point is to observe that $(h \circ f)' \in \mathrm{BV}_{loc}((m, b])$ implies that g' can be chosen to be finite at $f(b)$.

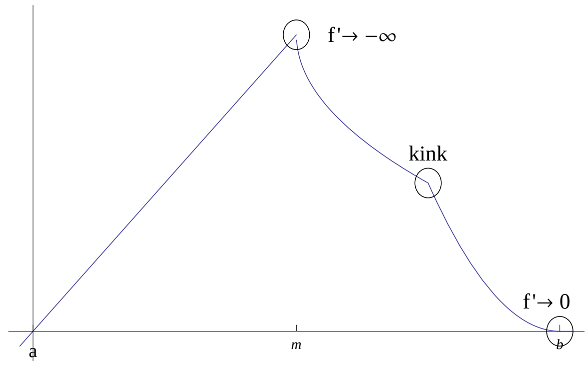

Figure 10: AFTER STRAIGHTENING ONE SIDE, CERTAIN PATHOLOGIES CAN EXIST ON THE OTHER SIDE AND STILL BE CONCAVIFIABLE.

Remark 9. We could have written Theorem 2's conditions (ii) and (iii) differently. For instance, f can be concavified if and only if after postcomposition by a function $h\colon [f(m),\infty) \to \mathbb{R}$, $\log \overline{D}(h\circ f)_1(x)$ and $\log \overline{D}(h\circ f)_2(x)$ are bounded above on $[0,t]$ for any $t < 1$, and bounded from below on $[s,t]$ for any $0 < s < t < 1$ and if the negative part of the resulting upper derivatives satisfy $[\overline{D}(\log \overline{D}(h\circ f)_i)]^- \in L^1_{loc}([0,1))$ for $i = 1, 2$. (Recall here that $f^-(x) = f(x)$ for $f(x) < 0$ and $f^-(x) = 0$ for $f(x) \geq 0$.) We have also used the fact that if $g \in L^1([a,b])$ then $G(x) = \int_a^x g(t)dt$ is absolutely continuous and in $\mathrm{BV}([a,b])$.

Remark 10. The necessity part of Theorem 2's proof can also be done using approximate derivatives. For each $\epsilon > 0$, let $r_\epsilon\colon \mathbb{R} \to \mathbb{R}$ be a positive smooth even unimodal function[9] with support $[-\epsilon,\epsilon]$ and $\int_{\mathbb{R}} r_\epsilon(t)dt = 1$ such that the nth derivative $r_\epsilon^{(n)}$ is an odd function for $n \in \mathbb{N}$ odd and an even function for $n \in \mathbb{N}$ even. We can use r_ϵ as a mollifier: starting from $f \in L^p$, for $p \in [1,\infty]$, the convolution f_ϵ defined by $f_\epsilon(x) = \int_{\mathbb{R}} r_\epsilon(x-t)f(t)\,dt$ is a smooth strictly increasing function which converges to f in L^p as $\epsilon \to 0$. If $f \in W^{k,p}$ then f_ϵ converges to f in $W^{k,p}$. Note that since f is strictly increasing and r'_ϵ is odd, $f'_\epsilon(x) = \int_{\mathbb{R}} r'_\epsilon(x-t)f(t)dt$ is strictly positive. In particular, f_ϵ is also strictly increasing.

In light of Remark **??**, with respect to Theorem 2's notation we need only check condition (iii). For $f\colon [a,b] \to \mathbb{R}$, we can always find for each $\epsilon > 0$, a

[9]Recall f is *even* if $f(-x) = f(x)$ and *odd* if $f(-x) = -f(x)$ for all x in the domain.

function $g_\epsilon \in C^2$ such that

$$(\log g'_\epsilon)' > (\log(h_\epsilon \circ f_{\epsilon}{\rfloor}_{(m,b]}^{-1})')',$$

where $h_\epsilon = -\left(f_{\epsilon}{\rfloor}_{(a,m]}\right)^{-1}$. The function g'_ϵ has a bounded limit if and only if, for all $\epsilon > 0$,

$$\lim_{\epsilon \to 0} \log(h_\epsilon \circ f_{\epsilon}{\rfloor}_{(m,b]}^{-1})' \in \mathrm{BV}_{loc}((f(m), f(b))). \tag{4.22}$$

Condition 4.22 is equivalent to Theorem 2's bounded variation condition (iii).

5 Objective Functions on $\mathbb{R}^n$ and More General Manifolds and Geodesic Metric Spaces

Our discussion so far has been for functions whose domain is $\mathbb{R}^1$. The next natural step would be to consider functions on $\mathbb{R}^n$. However, with a little extra effort we will make the leap to consider functions on an arbitrary geodesic metric space X with a distance function d: $X \times X \to \mathbb{R}$.[10] This is a vast generalization at low cost, allowing for spaces of innumerable types of behavior (allowing for infinite dimensions, fractal pathologies, graphs, surfaces, and so forth). Distance must be defined, but the space need not have a norm, so distance can be purely ordinal. That the space be geodesic does rule out disconnected spaces, e.g., a space consisting of two disjoint line segments. We do not exclude non-proper[11] geodesic metric spaces, e.g. the Banach space of all differentiable functions of one variable with the C^1 norm. Applications of infinite or arbitrary dimensional domains arise when considering consumption over an unspecified or infinite number of years, whether time is continuous or discrete, choice of contract functions over a space of functions, choice of present action given an arbitrary parameter space of histories, and maximization of profit by choice of network design for employees.

[10]Recall that a geodesic metric space is a space X such that there is a curve γ ("a geodesic") between any two points $x, y \in X$ such that the distance $d(x, y)$ is realized by the length of γ, also measured with respect to d (see e.g. Papadopoulos (2005)).

[11]A metric space X is proper if its closed metric balls $B(x, r) = \{y \in X : d(x, y) \leq r\}$ are compact.

We will also look at a special case of a geodesic metric space, namely the smooth Riemannian[12] manifold M, in which case we will discover that conditions for concavifiability can be found that are amenable to easy verification. Manifolds are objects that are locally like $\mathbb{R}^n$, e.g. planes, donuts, and spheres, so they include $\mathbb{R}^n$ as a special case. Economists ordinarily work in $\mathbb{R}^n$, but we will go beyond it here since the added complexity is not too great and the results may be of interest in special economic applications (e.g., function spaces) and to mathematicians. We allow M to be a manifold with boundary (e.g. a smooth subdomain of a larger manifold.) In all cases we will assume the regularity of the manifold: that its metric and its boundary are sufficient to support the regularity required of the functions. The required relative regularity is straightforward to determine in particular cases, but is slightly different for each case, so we will leave this to the reader.

We will now expand our definition of quasiconcavity to spaces more general than $\mathbb{R}^1$.

Definition 3. (QUASICONCAVITY AND CONCAVITY ON A GEODESIC METRIC SPACE) A function $f\colon X \to \mathbb{R}$ on a geodesic metric space is *quasiconcave* if and only if $f \circ \gamma\colon [0,1] \to \mathbb{R}$ is quasiconcave for every geodesic $\gamma\colon [0,1] \to \mathbb{R}$, where we assume that γ is parameterized proportional to, but not necessarily by, arclength.[13]

Similarly, f is *concave* if and only if for each geodesic $\gamma\colon [0,1] \to X$, $f \circ \gamma$ is concave as a function on $[0,1]$.

Since geodesics are straight lines in $\mathbb{R}^n$ with its standard metric, in Euclidean space this definition agrees with our definitions for $\mathbb{R}^n$ at the start of the paper.

In what follows, we let $m \in X$ be the unique point maximizing f if f is quasiconcave or minimizing f if $-f$ is quasiconcave. Also recall that for a function $F : X \to \mathbb{R}$, the negative part of F, F^-, is defined by

$$\boxed{F^-(x)} \equiv \begin{cases} F(x) & F(x) < 0 \\ 0 & F(x) \geq 0 \end{cases}.$$

[12]That is, M is equipped with a Riemannian metric coming from an inner product g_x on each tangent space T_xM.

[13]That is, we assume that our geodesics are parameterized so that if $\gamma(0) = x$ and $\gamma(1) = y$ then $d(\gamma(s), \gamma(t)) = |t - s|\, d(x, y)$ for all $s, t \in [0,1]$.

We can now state the complete criterion for concavification of quasiconcave functions, the generalization of Theorem 2 for geodesic metric spaces instead of R^1.

Theorem 3. *Let X be any geodesic metric space. For any continuous function $f\colon X \to \mathbb{R}$ there is a function $g\colon \mathrm{Range(f)} \to \mathbb{R}$ such that $g \circ f$ is strictly concave if and only if:*

(i) Either f or $-f$ is strictly quasiconcave;

(ii) $h \circ f$ is locally Lipschitz on $X - \{m\}$ where $h = (f \circ \gamma_o)^{-1}$ for some geodesic γ_o in X ending at m such that $\mathrm{Range}(f \circ \gamma_o) = \mathrm{Range}(f)$. Moreover $\overline{D}(h \circ f \circ \gamma)$ does not vanish on any geodesic segment $\gamma\colon (0,1) \to X$ for which $h \circ f \circ \gamma$ is strictly increasing.

(iii) The total variation of $\log \overline{D}(h \circ f)$ along all geodesics $\gamma\colon [0,1] \to X$ for which $h \circ f \circ \gamma$ is strictly increasing is uniformly bounded away from the extrema of $h \circ f$, or in other words,

$$\inf_{\gamma} \left\{ [\overline{D}(\log \overline{D}(h \circ f \circ \gamma)^{-1})]^{-} \right\} \in L^1_{loc}(R). \tag{5.1}$$

(Here R is the interior of the range of f and the infimum is taken over all geodesic segments $\gamma\colon [0,1] \to X$ for which $h \circ f \circ \gamma$ is strictly increasing.)

Proof. Suppose first that the conditions are met. Let $q = \inf_{\gamma} \left\{ [\overline{D}(\log \overline{D}(h \circ f \circ \gamma)^{-1})]^{-} \right\}$. From Theorem 2's proof, any function g_0 such that $(\log g_0')' < q$ will concavify $h \circ f \circ \gamma$ for each such γ. Choose g_0 such that $(\log g_0')' = -1 + q$. (Note that we can extend g_o to a function at the endpoints of R as well.)

Observe that every segment $\gamma\colon [0,1] \to X$ contains a subsegment $[s,1]$ where $h \circ f \circ \gamma$ is strictly increasing on $[s,1]$. Moreover, $h \circ f \circ \gamma(1-t)$ is also strictly increasing for $t \in [1-s, 1]$. We now note that a quasiconcave function that is concave on both its strictly increasing and strictly decreasing part separately is concave. Hence, the function g_0 concavifies $h \circ f \circ \gamma$ for every geodesic γ, and so $g_0 \circ h \circ f$ is concave. (This applies even to geodesics through

m since $g_0 \circ h \circ f$ is concave on every subinterval on either side of m.) Taking $g = g_0 \circ h$ finishes this direction of the proof.

Conversely, suppose that there is a function g such that $g \circ f$ is concave. Since h is invertible, we may write $g \circ f = g \circ h^{-1} \circ h \circ f$, and set $g_o = g \circ h^{-1}$. By concavity of $g \circ f$ along γ_o, we have that $g_o = g \circ f \circ \gamma_o$ is convex. In particular it C^1 with $\log(g_o)'$ Lipschitz with derivative belonging to $L^1_{loc}(R)$. Moreover, for any geodesic γ: $[0,1] \to X$, with $f \circ \gamma$ strictly increasing, we have $(\log(g_o)')' \leq q_\gamma$ where $q_\gamma = [\overline{D}(\log \overline{D}(h \circ f \circ \gamma)^{-1})]^-$, the comparison holding almost everywhere. Taking infima over all such γ implies $(\log(g_o)')' \leq q$ as desired.

$\square$

Put crudely, Theorem 3 says that in any geodesic metric space, including infinite- dimensional ones, the function f being strictly concavifiable by a strictly increasing g is equivalent to three conditions on f. First, f must be strictly quasiconcave. Second, after being straightened to linear along one geodesic spanning the whole range, the resulting function must not be too flat or too steep in any direction. Lastly, condition (iii) requires that the total variation of the log of the derivative along all geodesic segments must be bounded uniformly, away from the endpoints.

Remark 11. Even in this most general of our theorems, the proof is still constructive. The function g which concavifies f can be explicitly constructed following the proof, using only data provided by the hypotheses. Additionally, if X is a proper metric space, then one can postcompose again to flatten the peak of the function as was done in constructing the g from Theorem 2, if one so desires.

Remark 12. Once again, condition (ii) is unnecessary, strictly speaking, in that it follows from the satisfaction of condition (iii), since otherwise $\log \overline{D}(h \circ f \circ \gamma)$ would not be defined as a function. We include it for parallelism with the other theorems, and because otherwise the condition might appear hidden in condition (iii).

Remark 13. There are two reasons a function may not be in L^1_{loc}: (a) it is not measurable; or (b) it does not have bounded integral on some compacta. We only need condition (iii) to rule out problem (b). Even though the composition of measurable functions is not always measurable, problem (a) cannot arise here since the following operations always result in a measurable function: the composition of measurable functions by continuous functions, the real limit of measurable functions (e.g., the $\overline{D}$ operator), taking the supremum of measurable functions, and taking the negative part of a measurable function. Note that $h \circ f \circ \gamma$ is continuous and hence measurable, and therefore the function $\sup_\gamma \left\{ [\overline{D}(\log \overline{D}(h \circ f \circ \gamma))]^- \right\}$ is automatically measurable.

Theorem 3 applies quite generally, but its conditions, though sharp, are not always easy to verify. In the case of a Riemannian manifold M, however, we can search for conditions that take advantage of the global smooth structure on M. In the next section we will examine a relatively simple set of necessary and sufficient conditions which allow for much easier verification. In particular, they avoid the potentially difficult task of deciding whether or not the restriction of the function to each geodesic lies in the required function space. The price to be paid for this simplification is the limitation to twice-differentiable functions, for both the original function and the concavifier.

5.1 Differentiable Functions on the Manifold M (including $\mathbb{R}^n$) as a Special Case

Here we consider a C^2 Riemannian manifold M with its Riemannian connection ∇ (the natural Riemannian extension of the Euclidean gradient differential operator). We will make the assumption that the quasiconcave function $f \colon M \to \mathbb{R}$ is twice differentiable. Later, for our last theorem, we will weaken this to functions belonging only to the Sobolev space $W^{2,1}$ (i.e. possessing weak second derivatives).

Let $f \colon M \to \mathbb{R}$ be strictly quasiconcave. Suppose ∇f exists and does not vanish at a point x. In a neighborhood of x, choose an orthonormal basis $\{e_i(x)\}$ such that (i) $e_1 = \frac{\nabla f}{\|\nabla f\|}$, and (ii) $e_2, \ldots, e_n$, tangent to the level set of f through x, are a diagonal basis for the second fundamental form[14] of the level set $f^{-1}(f(x))$.

The Hessian of a function f on an arbitrary Riemannian manifold is the

[14]This is the symmetric form describing the shape operator of relative curvatures of the embedded manifold.

$(0, 2)$ tensor[15] $\text{Hess}(f) = \nabla df$. Given a basis, $e_1, \ldots, e_n$ at a point p, the corresponding matrix of the Hessian at p has entries

$$f_{ij} = \langle \nabla_{e_i}(\nabla f), e_j \rangle = \nabla_{e_i} \langle \nabla f, e_j \rangle - \langle \nabla f, \nabla_{e_i} e_j \rangle \tag{5.2}$$

This matrix depends on the metric, and not just on the smooth structure (except at critical points of the function f, where $\nabla df = d^2 f$). Note, too, that $\text{Hess}(f)$ is symmetric, which can be seen easily by extending the basis $\{e_i\}$ to a coordinate basis so that

$$f_{ij} - f_{ji} = e_i(df(e_j)) - e_j(df(e_i)) - df(\nabla_{e_i} e_j - \nabla_{e_j} e_i) = [e_i, e_j](f) - df([e_i, e_j]) = 0. \tag{5.3}$$

Thus equipped, we can present a necessary and sufficient condition for concavifiability of a twice-differentiable quasiconcave function f by postcomposing with a twice-differentiable function g. (Recall that q its negative part q^- is the negative part of q, as defined as in Equation (4.20).) Our theorem will apply to C^2 Riemannian manifolds; that is, manifolds admitting C^2 charts for which the metric tensor coefficients are also C^2 functions. It will depend on the principal curvatures of the level sets of f, which are the eigenvalues of the second fundamental form of a submanifold. These values (which are positive for a quasi-concave function) indicate the bending of the submanifold relative to the ambient manifold's curvature. For an $\mathbb{R}^2$ example to illustrate the theorem, look ahead to Figure 11s example after the proof.

[15] A $(0, q)$-tensor is, roughly speaking, a multilinear map that eats q distinct vectors and spits out a scalar. For $q = 2$, an example is the dot product in $\mathbb{R}^n$.

Theorem 4. *Let M be a C^2 Riemannian manifold. For any twice-differentiable function $f\colon X \to \mathbb{R}$ there is a strictly increasing twice-differentiable $g\colon \mathbb{R} \to \mathbb{R}$ such that $g \circ f$ is strictly concave if and only if:*

(i) f is strictly quasiconcave;

(ii) ∇f does not vanish except possibly at the maximum point of f, HOW ABOUT A HALF-DERIVATIVE AT AN EDGE OF THE RANGE e .g. $f(x) = x^2$ on [0,1]? And

WHERE IN THE PROOF DOES IT COME IN THAT ITS OK FOR THE GRADIENT TO DISAPPEAR AT A MAXIMUM?

(iii) The function q^- belongs to $L^1_{loc}(R)$, where R is the interior of the range of f and q is defined by

$$q(t) = \inf \frac{1}{\|\nabla f\|^2} \left(-f_{11} - \sum_{j=2}^{n} \frac{f_{1j}^2}{\lambda_j \|\nabla f\|} \right) \tag{5.4}$$

where the infimum is taken over points on the level set $f^{-1}(t)$ and where $\lambda_2, \ldots, \lambda_n$ denote the principal curvatures of the level sets of f.

A suitable g is

$$g(z) = \int_{f(m)}^{z} e^{\int_{f(m)}^{s} (-1+q_o(t))\, dt}\, ds, \tag{5.5}$$

with q_o a continuous function almost everywhere less than q^-.

Proof. Consider f in a neighborhood of a point $p \in M$. We assumed $\nabla_p f \neq 0$ which allows us to choose an orthonormal basis of $T_x M$[16] for x in a neighborhood of p as before, so that $e_1(x) = \frac{\nabla_x f}{\|\nabla_x f\|}$ and $e_2(x), \ldots, e_n(x)$ is a basis of $e_1^{\perp}$ which diagonalizes the second fundamental form of the level set $f^{-1}(f(p))$ at the point p. We will denote the eigenvalues, the principal curvatures for

[16] Here, $T_x M$ is the vector space of all tangent vectors to M at x.

this symmetric bilinear form, by $\lambda_2, \ldots, \lambda_n$, where $n = \dim(M)$. By quasi-concavity of f (condition (i)), these are all strictly positive. In terms of our basis we have $\lambda_j = -\left\langle \nabla_{e_j} e_1, e_j \right\rangle$ (see e.g. Chavel (1993)).

For any twice-differentiable function $g\colon \mathbb{R} \to \mathbb{R}$, the Hessian of $g \circ f$ is given by $\nabla^2(g \circ f) = (g'' \circ f) df \otimes df + g' \circ f \operatorname{Hess}(f)$. We need to show that this is negative definite under the hypotheses.

Recall that in the above frame we computed the (i, j)-entry of the Hessian to be $f_{ij} = \langle \nabla_{e_i}(\nabla f), e_j \rangle = \nabla_{e_i} \langle \nabla f, e_j \rangle - \langle \nabla f, \nabla_{e_i} e_j \rangle$. By our choice of frame, for $j > 1$ the term $\langle \nabla f, e_j \rangle$ identically vanishes, and so

$$f_{ij} = -\langle \nabla f, \nabla_{e_i} e_j \rangle = -\|\nabla f\| \langle e_1, \nabla_{e_i} e_j \rangle = -\|\nabla f\| \left(\nabla_{e_i} \langle e_1, e_j \rangle - \langle \nabla_{e_i} e_1, e_j \rangle\right) = \|\nabla f\| \left(\langle \nabla_{e_i} e_1, e_j \rangle\right).$$

In particular $f_{ii} = -\|\nabla f\| \lambda_i$ when $i > 1$. Putting this together we compute the Hessian of $g \circ f$ to be,

$$\nabla^2(g \circ f) = (g'' \circ f) \begin{pmatrix} \|\nabla f\|^2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} + (g' \circ f) \begin{pmatrix} f_{11} & f_{12} & f_{13} & \cdots & f_{1n} \\ f_{21} & -\lambda_2 \|\nabla f\| & 0 & \cdots & 0 \\ f_{31} & 0 & -\lambda_3 \|\nabla f\| & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f_{n1} & 0 & \cdots & 0 & -\lambda_n \|\nabla f\| \end{pmatrix} \tag{5.6}$$

Note we have $f_{1j} = -\langle \nabla f, \nabla_{e_1} e_j \rangle$, and moreover,

$$f_{11} = \frac{1}{\|\nabla f\|^2} \langle \nabla_{\nabla f} \nabla f, \nabla f \rangle = \frac{\nabla_{\nabla f} \|\nabla f\|^2}{2 \|\nabla f\|^2} = \frac{\nabla_{\nabla f}(\|\nabla f\|)}{\|\nabla f\|} = \partial_{e_1} \|\nabla f\|,$$

or, in other words, f_{11} is the growth rate of $\|\nabla f\|$ in the ∇f direction.

Similarly, since $\langle e_1, e_1 \rangle = 1$ identically, $\left\langle e_1, \nabla_{e_j} e_1 \right\rangle = \frac{1}{2} e_j(\langle e_1, e_1 \rangle) = 0$. Therefore,

$$f_{1,j} = f_{j,1} = \nabla_{e_j} \langle \nabla f, e_1 \rangle - \left\langle \nabla f, \nabla_{e_j} e_1 \right\rangle = \nabla_{e_j} \|\nabla f\|. \tag{5.7}$$

Since the values λ_i are all positive, we see that the principal minors, starting from the lower right, alternate sign. Hence in order to show that the eigenvalues of $\operatorname{Hess}(g \circ f)$ are all negative it remains to show that the sign of the entire determinant is $(-1)^n$.

Observe that for $j > 1$, the minor of the combined matrix corresponding to the pivot f_{1j} along the first row becomes lower triangular after moving the row whose entry begins with f_{j1} to the first row. This introduces a $(-1)^{j-2}$ to the determinant of the minor, which is then $(-1)^{j-2}(g' \circ f)^{n-1} f_{j1}\lambda_2 \ldots \lambda_{j-1}\lambda_{j+1} \ldots \lambda_n(-\|\nabla f\|)^{n-2}$. In particular the cofactor for $j > 1$, namely $(-1)^{j-1} f_{1j}(g' \circ f)$ times this, is simply

$$(-1)^{n-1}(g' \circ f)^n \|\nabla f\|^{n-2} \frac{f_{1j}^2}{\lambda_j} \prod_{i=2}^{n} \lambda_i.$$

Hence, adding in the first cofactor, the entire determinant of $\mathrm{Hess}(g \circ f)$, found by expanding on minors across the first row, yields

$$\det \mathrm{Hess}(g \circ f) = (-1)^{n-1} \left(\frac{g'' \circ f}{g' \circ f} \|\nabla f\|^2 + f_{11} + \sum_{j=2}^{n} \frac{f_{1j}^2}{\lambda_j \|\nabla f\|} \right) (\|\nabla f\|)^{n-1} (g' \circ f)^n \prod_{i=2}^{n} \lambda_i. \tag{5.8}$$

We care about when expression (5.8) has sign $(-1)^n$. Since $\|\nabla f\|$, g', f_{1j}^2 and λ_i are all positive, this happens if and only if

$$(g'' \circ f) \leq \frac{(g' \circ f)}{\|\nabla f\|^2} \left(-f_{11} - \sum_{j=2}^{n} \frac{f_{1j}^2}{\lambda_j} \right), \tag{5.9}$$

which can be satisfied for a given f by a g with $g' > 0$ and $g'' < 0$, provided that for almost every value y in the range of f, the quantity

$$\frac{1}{\|\nabla f\|^2} \left(-f_{11} + \sum_{j=2}^{n} \frac{f_{1j}^2}{\lambda_j \|\nabla f\|} \right)$$

is bounded below by $q(y)$ on the level set $f^{-1}(y)$, where $q \in L^1_{loc}$. By the theorem's assumption we have such a bound. The g function in the statement of Theorem 4 satisfies these necessary conditions.

Conversely, ∇f must be bounded, away from the maximum point m, by Theorem 3's condition, and if $q \notin L^1_{loc}$ then we cannot obtain a g which everywhere satisfies the needed inequality.

□

Theorem 4 generalizes the "one-point" conditions of Fenchel (1953) for $\mathbb{R}^n$ (as reformulated in Section 4 of Kannai (1977)) to the Riemannian setting. Kannai's condition (I) on utility v corresponds to our condition (ii) on f. However he is allowing for weak concavifiability, which accounts for his necessary conditions (II) and (III) differing from our condition (i) when the sublevel sets of v are not strictly convex. Otherwise, these conditions are equivalent to our condition (i) and his conditions (IV) and (V) are folded into our condition (iii). This is best seen through the rephrasing of Kannai's condition (IV) as (IV′) and noting that his k equals our $\|\nabla f\|$ and that under our assumptions in the case when $M = \mathbb{R}^n$, $-\lambda_j \|\nabla f\| = f_{jj}$ under our assumptions when $M = \mathbb{R}^n$. Note that Kannai's perspective is that of constructing a concave utility function based on weakly convex preference relations, whereas we start with an arbitrary function and see if it can be concavified.

Figure 11s example: What condition (iii) excludes. Condition (iii) can be easily violated by a C^2 function f satisfying conditions (i) and (ii) by making the (necessarily noncompact) level sets become asymptotically flat sufficiently quickly as points tend to infinity. A simple example is the quasiconcave function $f(x,y) = e^{e^x} y$ defined in the open positive quadrant of $\mathbb{R}^2$, shown in Figure 11. Its gradient, $\left(e^{x+e^x} y, e^{e^x}\right)$, is nonvanishing and its Hessian restricted to the level set of value t as a function of the x coordinate is $-\frac{te^x(e^x-1)}{t^2 e^{2x-2e^x}+1}$. The negative definiteness shows that f is strictly quasiconcave. The quantity in condition (iii) on the level set of t works out to be $\frac{1}{-t+te^{-x}}$, whose infimum is always $-\infty$, and thus f is not concavifiable.

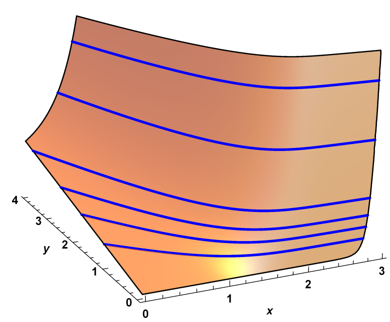

Figure 11: A FUNCTION VIOLATING CONDITION (III) OF THEOREM 4

Remark 14. Since $f_{1j} = -\langle \nabla f, \nabla_{e_1} e_j \rangle$, in the special case that the integral curves of the vector field ∇f lie along geodesics of M, then $f_{1j} = 0$ for all $j > 1$. This occurs, for instance, when f is constant on distance spheres about a fixed point.

Remark 15. Some twice-differentiable functions f with ∇f vanishing at points other than the maximum can also be concavified, provided we are willing to concavify using a g which is not twice differentiable. The more general condition is that after an initial postcomposition by a non-twice differentiable function g_o the resulting $g_o \circ f$ must satisfy conditions (ii) and (iii). In particular, when ∇f vanishes at a point, it must do so on the entire level set, though this alone is not sufficient.

Remark 16. In contrast to Theorems 2 and 3, here f is Lipschitz from the beginning, by virtue of being twice differentiable, and moreover $\log(h \circ f \circ \gamma)'$ automatically belongs to BV_{loc} for any twice differentiable increasing function h and geodesic γ under the assumption of condition (ii). Also, Theorem 4's condition (iii) is vacuous for one-dimensional M and C^2 function f when condition (ii) holds. So testing the theorem with one-dimensional examples is pointless.

If f is C^2 with nonvanishing gradient, then the quantity (5.4) in the infimum of the definition of $q(t)$ in condition (iii) of Theorem 4 is uniformly bounded and continuous on compact sets. Moreover, the infimum of any compact family of continuous functions is always continuous. Hence, we immediately obtain that the variation function q from (5.4) is continuous if f is C^2 with compact level sets. We express this as the following especially simple corollary.

Corollary 1. *If $f\colon M \to \mathbb{R}$ is strictly quasiconcave and C^2, with compact level sets, then there is a C^2 strictly concavifying g if and only if ∇f does not vanish except at f's maximum.*

Remark 17. Fenchel's Example in Figure 4 does not satisfy Corollary 1s conditions, because it is not strictly quasiconcave. In fact, for any function not strictly quasiconcave, at least one of the principal curvatures λ_i vanishes somewhere and thus quantity (5.4) becomes unbounded.

We can also work with the weak Hessian of f for functions $f \in W^{2,1}$. A weak gradient for $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is any vector function $\phi\colon \Omega \to \mathbb{R}^n$ such that for every smooth compactly supported function $\rho\colon \Omega \to \mathbb{R}$,

$$\int_\Omega \phi(x)\rho(x)dx = -\int_\Omega f(x)(\nabla\rho)(x)dx. \tag{5.10}$$

We will denote any such weak gradient by ∇f. This is justified, since from the definition any two weak gradients agree almost everywhere. By taking charts and using the volume forms one can see that this definition extends to arbitrary tensors on arbitrary smooth manifolds.[17] Define $W^{0,p}(M)$ to be $L^p(M)$, which we extend to denote the space of L^p tensors of any type on M. We then extend inductively by defining $f \in W^{k,p}$ if $\nabla f \in W^{k-1,p}$. We only used L^1 existence of second derivatives in Theorem 4's proof, so we obtain another corollary.

Corollary 2. *The results of Theorem 4 and Corollary 1 hold verbatim for* $f \in W^{2,1}$.

For $f \in W^{2,1}$, the smooth local convolution functions $f_\epsilon \colon M \to \mathbb{R}$ have Hessians which converge in L^1 to the weak Hessian of f. Hence, the $W^{2,1}$ Sobolev space provides a class of functions which are reasonably easy to work with in the sense that the condition in Theorem 4 is "differential" and hence easy to check. The space $W^{2,1}$ is a fairly large, and flexible, class frequently used in the theory of partial differential equations because of its closure properties and techniques for embedding it in other function spaces.

6 Discontinuous or Weakly Quasiconcave Objective Functions

We did not require differentiability for Theorems 2 and 3, but we did assume functions were continuous. Also, we characterized strict quasiconcavity, but not quasiconcavity generally. It turns out that our results are true for some but not all discontinuous functions and that we can also characterize quasiconcavity generally.

6.1 Discontinuous Functions

First, let us see why our theorems are false for discontinuous functions generally. For simplicity we will assume throughout the section that we are in the case where the domain is a manifold M.

By collapsing intervals in the range of a quasiconcave $f \colon M \to \mathbb{R}$ we can always find a nondecreasing g such that $g \circ f$ is continuous. However, g might

[17] A chart is a bijective continuous mapping of an open set of a manifold onto $\mathbb{R}^n$.

have to be the constant function since the interval gaps between values at discontinuities can cover the entire range of f.

Figure 12s example: A discontinuous quasiconcave function. The following strictly quasiconcave function f: $[0,1] \to \mathbb{R}$ is shown in Figure 12:

$$f(x) = \begin{cases} 5 + 10x & x \leq \frac{1}{2} \\ 5 - 10(x - .5) & x > \frac{1}{2}. \end{cases} \tag{6.1}$$

The dotted line shows a particular $g(f(x))$. It must have a discontinuous drop at $x = .5$ because for x less than 5, $g(5) > g(x)$, a strict inequality. Such a discontinuity prevents $g(f(x))$ from being concave.

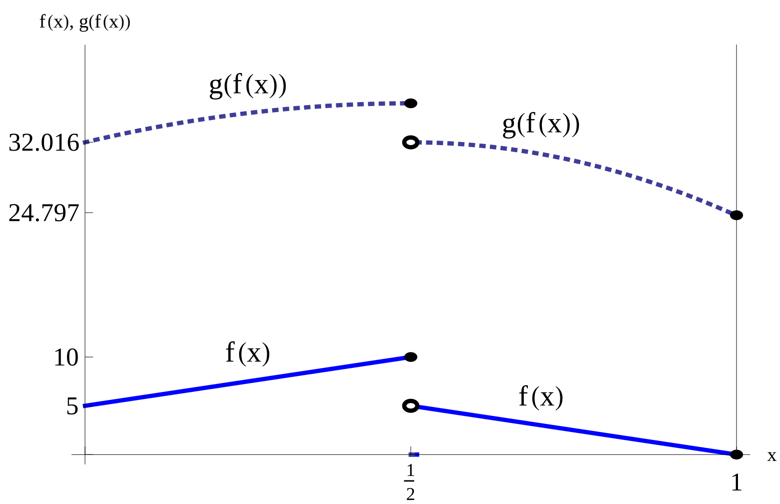

Figure 12: A Discontinuous Quasiconcave Function Which Cannot Be Mapped to a Concave Function

What we would like to do is to postcompose f by a function h to make it continuous, so we could apply our earlier sections' theorems. We can do that for certain discontinuous functions, such as the $f(x)$ in Figure 13.

Recalling that $f(x)$ is defined over (a, b) and reaches its maximum at m, define $A \equiv (a, m)$ and $B \equiv [m, b)$ (see Figure 13). Define R_A as the union of the images of $f(x)$ from A and R_B as the union from B. Define the gap sets

$$G_A \equiv [\inf_{x \in A} f(x), f(m)) - R_A \quad \text{and} \quad G_B \equiv [\inf_{x \in B} f(x), f(m)] - R_B.$$

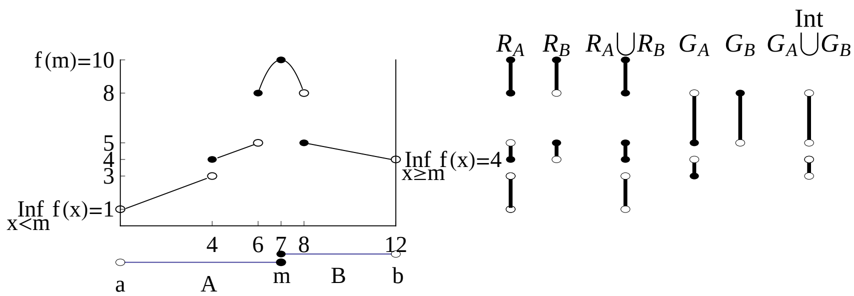

Figure 13: A Concavifiable Discontinuous Function

> **Proposition 1.** *For any, possibly discontinuous, strictly quasiconcave $f\colon M \to \mathbb{R}$, there is a function $h\colon f(M) \to \mathbb{R}$ such that $h \circ f$ is strictly quasiconcave and continuous if and only if the range of f, $R_A \cup R_B$, and the interior of the union of the gap sets, Interior $G_A \cup G_B$, are disjoint. Any such h may be chosen to be nondecreasing.*

Proof. If the hypothesis applies, then whenever $s > t$, the (closed) superlevel sets satisfy either $S_t(f) = S_s(f)$ or else $S_t(f)$ is contained in the interior of $S_s(f)$.

The hypothesis is exactly the hypothesis that the discontinuous jumps in f occur along entire level sets and are the same height. So each gap in the image can be closed by a piecewise linear nondecreasing h which is constant on the gap interval. The gaps are disjoint so the whole procedure can be accomplished by a single h. Since no interval of level sets are sent to the same value, strict quasiconcavity is clearly preserved by this operation.

Conversely, suppose h makes $h \circ f$ continuous and strictly quasiconcave but the hypothesis on the gaps of f is not satisfied. Let $[a, b] \subset \text{Range}(f)$ be an interval in the range of f where for each $t \in [a, b]$ the superlevel sets $S_t(f)$ share a common boundary point. Then h must carry these to the same value. However, by assumption they are not the same superlevel set, so there is a nondegenerate curve $c\colon [a, b] \to M$, and so $h \circ f$ is constant on $c(t)$. This contradicts strict quasiconcavity.

□

7 Weakly Quasiconcave Objective Functions

One might think that if strictly quasiconcave functions satisfying Theorem 3's necessary conditions are those that can be transformed strictly monotonically into concave functions, then we should have a similar statement for weakly quasiconcave functions, along the lines of the following.

Conjecture 1. *If weakly quasiconcave functions satisfy an analogue to Theorem 3's conditions, then they can be transformed weakly monotonically into weakly concave functions.*

Unfortunately, this conjecture is false.

Figure 14s example: A weakly quasiconcave function. Figure 14 shows a function f which is only weakly quasiconcave because the values of x between 1 and 2 map to the same $f(x)$ (and likewise for values between 3 and 5). It is still quasiconcave because the upper level set $[1, 5.5]$ is convex and though multiple values of x maximize $f(x)$, the only maximal value is 2.

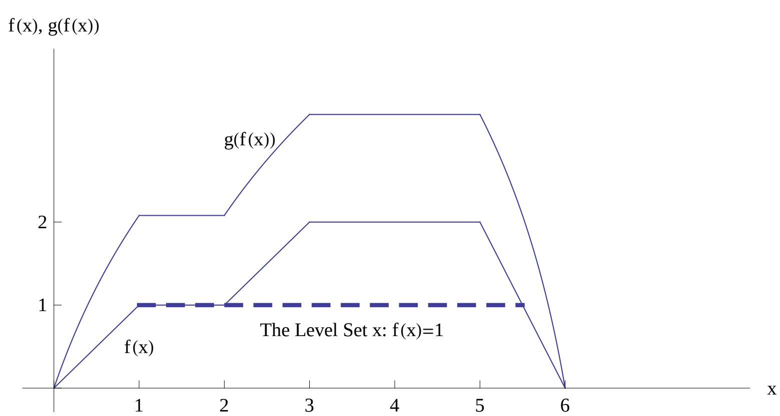

Figure 14: A Weakly Quasiconcave Function

If we apply a monotonically increasing g to the function f in Figure 14, that g will have to map all the values of x in $[1, 2]$ to the same $g(f(x))$. As a result, $g(f(x))$ will have a flat part as shown and cannot be even weakly concave.

The conjecture is false because the word "weakly" is used differently for concavity and for quasiconcavity. Consider a twice-differentiable function

$f\colon \mathbb{R} \to \mathbb{R}$. If f is strictly concave, it cannot have any linear portions; its second derivatives cannot vanish (though they can if it is only weakly concave, e.g., $f(x) = x$). If f is strictly quasiconcave, it can have linear portions. The function f is only weakly quasiconcave if it has horizontal segments aside from its peak; that is, if the first derivative vanishes over some segment other than the maximum. The slope's rate of change is irrelevant for quasiconcavity. What does matter about the slope— the essence of quasiconcavity— is that it must not switch sign more than once. What make f weakly quasiconcave is if its slope is zero over an interval, coming as close as possible to switching sign more than once.

What can we do, then, to extend our constructions to concavify weakly concave functions? What we will do is approximate the function f by a series of strictly concavifiable functions. Recall that even strictly quasiconcave functions may not be concavifiable, as not all such functions satisfy Theorem 3's regularity conditions. Therefore we would also like to make sure that our strictly quasiconcave approximations are concavifiable. Fortunately, both properties of approximations are easy to arrange with the "connect-the-dots" approach we use below. For simplicity we will assume that the domain of f is a convex subset $D \subset \mathbb{R}^n$, but the constructions and resulting properties apply straightforwardly to convex domains in any Riemannian manifold since they are purely local and Riemannian manifolds always have small convex neighborhoods around each point which look similar to a convex set in $\mathbb{R}^n$.

We will use a two-step process. The first step is to approximate the function by one whose level sets all have empty interior.

At most a countable number of level sets for f can contain nonempty interior, since any disjoint family of open sets in $\mathbb{R}^n$ is countable, and the same holds for manifolds. We may enumerate these exceptional level sets by $f^{-1}(c_1), f^{-1}(c_2), \ldots$, where the values c_i are not necessarily in increasing order. For now, fix a choice of small $\epsilon > 0$. Also, fix the constant $k_1 = 2^{-1}$. We make a new function f_1 from f as follows. For all values of $c \leq c_1$ that belong to the range of f, we set $f_1^{-1}(c - \epsilon k_1)$ to be the set at distance ϵk_1 from the super-level set $S_c(f)$. Whenever there are points other than $S_c(f)$ contained in the set bounded by $f_1^{-1}(c - \epsilon k_1)$, we assign the f_1 value of $c - \epsilon 2^{-1}$ to these as well. For values $c \leq c_1$ not in the Range(f) we exclude the value $c - \epsilon k_1$ from Range(f_1). For all $c > c_1$ and $x \in S_c(f)$, we set $f_1(x) = f(x)$. We finish by explaining how to assign the f_1 values in the range interval $(c_1 - \epsilon k_1, c_1]$. Let $A = \cup_{c > c_1} S_c(f)$, i.e. the open c_1 super-level set of f, and set B to be the set of all points at distance at most ϵk_1 from $S_{c_1}(f)$. Since A is weakly convex, the integral curves tangent to the gradient vector field of the function $x \mapsto d(x, A)$, where $d(x, A)$ is the distance from x to the set A, consist of geodesic segments. For any point $x \in B \setminus A$, let $l(x) \geq \epsilon k_1$ be the length of the maximal such integral geodesic segment passing through x and contained in $B \setminus A$. We then set $f_1(x) = c_1 - d(x, A)\frac{\epsilon k_1}{l(x)}$. This defines a new function f_1 whose domain consists of all points at distance at most ϵk_1 of the domain D of f.

Since uniform distance sets to weakly convex sets are weakly convex, the super-level sets of f_1 remain weakly convex, and so f_1 is again weakly quasiconcave. We now construct $f_2, f_3, \dots$ successively in a similar manner. Namely, for $i \geq 2$, f_i is constructed from f_{i-1}, by replacing f in the construction above by f_{i-1}, the constant $k_1 = 2^{-1}$ everywhere by the constant $k_i = 2^{-i}$, and c_1 everywhere by $c_i - \epsilon \sum_{j=1}^{i-1} k_j$. After repeating this process a countable number of times, we arrive at the limiting function, denoted $f_\epsilon = \lim_{i\to\infty} f_i$. This function inherits certain nice properties.

Lemma 5. *The function f_ϵ has the following properties:*

(i) f_ϵ is weakly quasiconcave,

(ii) f_ϵ has level sets with empty interior, and

(iii) f_ϵ converges pointwise almost everywhere to f as $\epsilon \to 0$, and it converges uniformly on compacta if f is continuous.

Proof. The first item was explained in the construction. For the second item, by the construction, all of the level sets with nonempty interior of f have been enlarged and broken into a new family of level sets with empty interior in f_ϵ.

Lastly, we note that the c-level set of f_ϵ is displaced at most by $\sum_{i=1}^{\infty} \epsilon k_i = \epsilon$ from that of f. Hence for any point x where a discontinuity of f does not occur, $f_\epsilon(x)$ converges pointwise to $f(x)$. There can be at most a countable set of level sets where a discontinuity occurs since f is quasiconcave. On each such level set, a discontinuity cannot occur at an interior point. Since the super-level sets are convex, and a countable union of measure 0 sets has measure 0, this is a measure 0 set of points.

If f is continuous then the displacement of the level sets implies that the convergence is uniform on compacta, although not necessarily globally uniform even on a single level set. $\square$

Remark 18. Note that f_ϵ is continuous when f is, but may not be if f is discontinuous. Even so, f_ϵ does not always converge pointwise at points of discontinuity of f. For instance, this occurs for functions that take on two values such as $f\colon [-1,1] \to \{0,1\}$ with $f^{-1}(0) = [-1,-\frac{1}{2}] \cup [\frac{1}{2},1]$ and $f^{-1}(1) = (-\frac{1}{2},\frac{1}{2})$ where the convergence does not occur at $x \in \left\{-\frac{1}{2},\frac{1}{2}\right\}$.

The second stage of the construction is to turn a weakly quasiconcave function f with the properties of f_ϵ above into a strictly quasiconcave function that is concavifiable.

Definition 4. (PIECEWISE LINEAR APPROXIMATION) Suppose $f\colon D \to \mathbb{R}$ is any weakly quasiconcave function, possibly discontinuous, whose level sets have nonempty interior. A **piecewise linear approximation** [18] f_ϵ is constructed for f as follows. Choose $\Lambda = \Lambda_\epsilon \subset D$ be a subset of the domain of f with the following properties:

1. Λ is a discrete set, that is it intersects any compacta in a finite set,
2. No two points in Λ belong to the same level set, i.e. they all have distinct values,
3. Each open ball $B(p,\epsilon)$ of radius ϵ in D contains at least one point of Λ.

We now produce a Delaunay-Voronoi triangulation [19] T corresponding to the discrete set Λ. The function f_ϵ is now defined to be the linear interpolation along each simplex of T of the values of f on the vertices of the simplex, which form a set of $n+1$ points belonging to Λ.

It is not immediately obvious that the above definition is well defined,

[18]This definition is quite different than the notion of "PL" map commonly used in mathematics. We use a very constrained triangulation, but we do not require links of points to be spheres, or even for the triangulation T to be a PL approximation of D.

[19]This is a simplicial complex T of a set of points $P \subset \mathbb{R}^n$ such that no point of P lies in the interior of the circumscribing ball of the vertices of any simplex of T. The triangulation T is unique if there is no k-dimensional plane containing $k+2$ points or k-sphere containing $k+3$ points for $1 \le k \le n-1$. For intuition, see Paul Chew's Java applet at http://www.cs.cornell.edu/info/people/chew/Delaunay.html.

particularly that such a set Λ exists. However, to satisfy the third criterion, any such ball intersects an uncountable continuum of level sets (otherwise the intersection with one of the level sets is open), and so a point can be always added to Λ (which is a countable set) within the ball on a new level set. This linear interpolation of f on each simplex is accomplished by standard barycentric averaging. Namely, if the vertices of a given simplex in T are $v_0, \ldots, v_n$ and $x = t_0 v_0 + \cdots + t_n v_n$ for any $t_0, \ldots, t_1 \in [0, 1]$ then the value at the interpolated point x is given explicitly by $f_\epsilon(x) = \sum_{i=0}^{n} t_i f(v_i)$. Lastly, since the Delauney-Voronoi triangulation of a discrete set in $\mathbb{R}^n$ always exists and consists of non-overlapping simplices (except on the boundaries where the definition of f_ϵ agrees), f_ϵ is well defined. When the domain D is a weakly convex subset of a manifold, the same is true provided that ϵ is chosen sufficiently small, so that all simplices lie inside a convex neighborhood of their vertices. Such neighborhoods are uniquely geodesic, and combinatorially behave like $\mathbb{R}^n$.

Proposition 2. *Suppose f is a weakly quasiconcave function whose level sets have no interior.*

(i) Every choice of piecewise linear approximation f_ϵ is strictly quasiconcave, continuous and strictly concavifiable, regardless of the regularity of f (e.g. the discontinuity or non-differentiability).

(ii) The functions f_ϵ converge to f pointwise almost everywhere and uniformly on compact domains if f is continuous.

Proof. Since the points of Λ belong to distinct level sets of f the values at the vertices are distinct and hence the linear interpolation has no geodesic line segments[20] on which f_ϵ is constant. Moreover, since the complex T is connected, the function f_ϵ is automatically continuous.

To prove concavifiability, we cannot simply apply the finite point domain version of Theorem 1 since we are interpolating, but it is still quite straightforward. We note that, by the construction, along any line segment in the domain there are only a finite number of piecewise linear changes, and no

[20] Or geodesics, in more general metric spaces. We have limited the proposition to $\mathbb{R}^n$ only for simplicity, but its extension is straightforward.

horizontal flat portions. After inverting one side, say f_1, of the resulting restricted function, the other side remains piecewise linear with a finite number of pieces, none of which are flat. Hence, the upper derivative along the resulting monotone increasing piece, say f_2, is always strictly positive, taking on only a finite number of values. The log of this function trivially belongs to BV_{loc} on its domain interval. In particular, it is concavifiable, by Theorem 3. We will prove the converse of statement (i) at the end of the proof.

For statement (ii), since the functions are locally monotonic we note that on any simplex of diameter at most ϵ in the domain, the difference $|f - f_\epsilon|$ is bounded by the supremum of the derivative of f_ϵ, not of f, on the simplex times at most ϵ. On each compact subset of the domain this uniformly converges to 0 as $\epsilon \to 0$. Hence, the function converges uniformly on compact subsets of the interior of D. The pointwise statement follows immediately. Note that the domain of f_ϵ is the realization of the simplicial complex T. This belongs to D by convexity of D, but may be a proper subset. Nevertheless these domains converge to D by the same reasoning above.

Lastly we note that for any x where f is continuous, then $\lim_{\epsilon\to 0} f_\epsilon(x) = f(x)$ regardless of the choice of f_ϵ, since the simplices containing x will eventually have vertices in Λ sufficiently near x, and thus their values also converge. As in Lemma 5s proof, the set of discontinuous points has measure 0. For continuous f the difference of the values of vertices of a simplex in T will depend uniformly on ϵ on compacta, by uniform continuity of f there. Hence f_ϵ converges uniformly on compacta to f in this case as $\epsilon \to 0$. In particular, it approaches f in any norm which relies only on pointwise differences (as opposed to derivatives, for instance). This includes any L^p norm for $p \in [1, \infty]$. $\square$

Proposition 2 is nice for practical purposes since Theorem 3's potentially arduous testing is unnecessary, nor even the more convenient testing of Theorem 4, provided one is willing to perturb the initial quasiconcave function as prescribed by Definition 4. Note, too, that in this proposition we have finally extended the idea of concavification to discontinuous functions. If we start with a general weakly quasiconcave function and first apply Lemma 5 followed by the above proposition, then we immediately obtain the following corollary.

Corollary 3. *Suppose f is any weakly quasiconcave function (possibly discontinuous). There exists a sequence of continuous strictly quasiconcave and* *concavifiable functions f_ϵ which converge to f as $\epsilon \to 0$ pointwise almost everywhere, and uniformly on compacta if f is continuous.*

8 Concluding Remarks

We have tried in this article to clarify the relationship between concavity and quasiconcavity at two levels. The first level is the intuitive one, where we have explored when continuous strictly quasiconcave functions, whether differentiable or not, are functions that can be monotonically blown up into strictly concave functions. The second level is the rigorous one, where we have shown that the intuition holds true for arbitrary geodesic metric spaces of any dimension, but subject to caveats that apply even in $\mathbb{R}^1$, caveats involving the Lipschitz continuity of the function and the amount of variation in its derivatives. We also show that these caveats become much simpler to check in the case of twice-differentiable functions with compact level sets, when all that is required is that the gradient of the function not vanish. To mediate between the intuitive and the rigorous we have provided numerous examples of quasiconcave functions that can and cannot be concavified.

In economics, the maximands most commonly encountered are utility functions, profit functions, and the payoff functions of principals and agents. Since calculus is our most powerful tool for maximization, we would like for these functions to be twice differentiable and concave. It would be nice if our results in this paper implied that whenever an objective function is quasiconcave, the modeller can transform it one-to-one to a twice-differentiable concave function. Alas, that is not true. We have shown that even if the quasiconcave objective function is nondifferentiable it can be transformed to a concave function, but that concave function is not necessarily differentiable, although it often can be made so. On the other hand, even nondifferentiable concave functions are always Lipschitz-continuous. This means that certain numerical optimization techniques are available which could not be used with the original, non-concave, objective function ("convex minimization" techniques— see Wikipedia [2011] and Boyd & Vandenberghe [2004]). Note, too, that the f_ϵ approximation that makes weakly quasiconcave functions strict can be chosen to not alter the size or place of the maximum and hence would be a suitable first step for numerical optimization. Whether the combination of monotonic and approximation transformations with convex minimization techniques would be useful in practical applications we do not know. However, we also have provided (in Theorem 4) an explicit function that concavifies, to a twice-differentiable $g(f(\cdot))$, any twice-differentiable strictly quasiconcave function f on bounded domains in $\mathbb{R}^n$ whose gradient does not vanish apart from the maximum.

References

Afriat, S. N. (1967) "The Construction of Utility Functions from Expenditure Data," *International Economic Review*, 8: 67–77.

Arrow, Kenneth J. & Alain C. Enthoven (1961) "Quasiconcave Programming," *Econometrica*, 29: 779–800.

Aubin, Thierry (1982) *Nonlinear Analysis on Manifolds. Monge-Ampre equations*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 252, Berlin, New York: Springer-Verlag (1982).

Aumann, Robert (1975) "Values of Markets with a Continuum of Traders," *Econometrica*, 43: 611–646.

Boyd, Stephen & Lieven Vandenberghe (2004) *Convex Optimization*. Cambridge University Press, free access at http://www.stanford.edu/~boyd/cvxbook/.

Chavel, Isaac (1993) *Riemannian Geometry— A Modern Introduction*, Cambridge Tracts in Mathematics, vol. 108, Cambridge University Press, Cambridge (1993).

DeFinetti, B. (1949) "Sulle Stratificazioni Convesse," *Annali di Matematica*, 30: 173–183.

Fenchel, W. (1953) "Convex Cones, Sets, and Functions," mimeo book, Princeton Mathematics Department, Princeton, New Jersey. Posted on the web with free access by Abros Press, http://rasmusen.org/x/abros/fenchel.pdf.

Folland, G. B. (1984) *Real Analysis. Modern Techniques and Their Applications*, Pure and Applied Mathematics Series, New York: John Wiley & Sons Inc.

Ginsberg, W. (1973) "Concavity and Quasiconcavity in Economics," *Journal of Economic Theory*, 6: 596–605.

Green, Jerry R., Andreu Mas-Colell & Michael D. Whinston (1995) *Microeconomic Theory*, Oxford University Press (1995).

Guerraggio, Angelo & Elena Molho (2004) "The Origins of Quasiconcavity: A Development between Mathematics and Economics," *Historia Mathematica*, 31: 62–75.

Hjertstrand, Per (2011) "A Generalized Afriat's Theorem," working paper, Department of Economics, Lund University, (February 2011).

Kannai, Yakar (1977) "Concavifiability and Constructions of Concave Utility Functions," *Journal of Mathematical Economics*, 4: l–56.

Kannai, Yakar (1981) "Concave Utility Functions— Existence, Construction, and Cardinality," pp. 543-612 of *Generalized Concavity in Optimization and Economics,* ed. by Siegried Schaible and William T. Ziemba, New York: Academic Press, 1981.

Kannai, Yakar (2005) "Remarks Concerning Concave Utility Functions on Finite Sets," *Journal of Economic Theory* 26: 333–344.

Kreps, David (1990) *A Course in Microeconomic Theory*, Princeton: Princeton University Press (1990).

Matzkin, Rosa L. & Marcel K. Richter (1991) "Testing Strictly Concave Rationality," *Journal of Economic Theory*, 53: 287–303.

Osborne, Martin (undated, untitled) U. of Toronto Economics Department notes, http://www.economics.utoronto.ca/osborne/MathTutorial/QCC.HTM.

Papadopoulos, Athanase (2005) *Metric Spaces, Convexity and Nonpositive Curvature*, IRMA Lectures in Mathematics and Theoretical Physics, vol. 6, European Mathematical Society (EMS), Zürich (2005).

Pogany, Peter (1999) "An Overview of Quasiconcavity and Its Applications in Economics," Office of Economics working paper, U.S. International Trade Commission, http://www.usitc.gov/publications/docs/pubs/research_working_papers/ec9904a.pdf, (April 1999).

Rapcsak, T. (2005) "Survey of the Fenchel Problem of Level Sets," in *Variational Analysis and Applications,* F. Giannessi and A. Maugeri (eds), pp. 935–956, Dordrecht, Holland: Kluwer Academic Publishers (2005).

Reny, Philip J. (2010) "A Simple Proof of the Nonconcavifiability of Functions with Linear Not-All-Parallel Contour Sets," working paper, Department of Economics, University of Chicago (October 2010).

Richter, Marcel K. & K.-C. Kam-Chau Wong (2004) "Concave Utility on Finite Sets," *Journal of Economic Theory,* 115: 341–357.

Takayama, Akira (1985) *Mathematical Economics*, 2nd edition (1st edition 1974), Cambridge: Cambridge University Press (1985).

Varian, Hal R. (1982) “The Nonparametric Approach to Demand Analysis,” *Econometrica*, 50: 945–973.

Wikipedia (2011) “Convex Optimization,” http://en.wikipedia.org/wiki/Convex_optimization or http://rasmusen.org/papers/quasi/Convex_optimization.pdf (May 24, 2011).

Wikipedia (2011) “Sobolev Inequality,” http://en.wikipedia.org/wiki/Sobolev_inequality.htm or http://rasmusen.org/papers/quasi/Sobolev_inequality.htm (June 18, 2009).

Wilson, Charles (2009) “Convex Sets, Concave and Quasiconcave Functions,” New York University notes, http://homepages.nyu.edu/~caw1/ (April 1, 2009).

Ziemer, William (1989) *Weakly Differentiable Functions.* Graduate Texts in Mathematics, New York: Springer-Verlag (1989).