TECHNICAL RESEARCH REPORT

Institute for Systems Research

On an Elementary Characterization of the Increasing Convex Ordering, with an Application

by A.M. Makowski

The Institute for Systems Research is supported by the National Science Foundation Engineering Research Center Program (NSFD CD 8803012), Industry and the University

TR 93-9

ON AN ELEMENTARY CHARACTERIZATION OF THE INCREASING CONVEX ORDERING, WITH AN APPLICATION

by

Armand M. Makowski[1]

ABSTRACT

In this short note, we present a simple characterization of the increasing convex ordering $\leq_{icx}$ on the set of probability distributions on $\mathbb{R}$. We show its usefulness by providing a very short proof of a comparison result for $M|GI|1$ queues due to Daley and Rolski and obtained by completely different means.

[1] Electrical Engineering Department and Institute for Systems Research, University of Maryland, College Park, MD 20742. The work of this author was supported partially through NSF Grants NSFD CDR-88-03012 and ASC-88-02764

1. INTRODUCTION

In this short note we discuss two simple yet useful characterizations of the increasing convex ordering $\leq_{icx}$ on the set of probability distributions on $\mathbb{R}$, and show its usefulness by providing a very short proof of a comparison result for $M|GI|1$ queues recently obtained by Daley and Rolski [5]. Before presenting the results in Theorem 1 below, we briefly recall some of the notions discussed here, with the notation being essentially that of [1]: Throughout, we find it convenient to define all the random variables (rvs) of interest on some common probability triple $(\Omega, \mathcal{F}, \mathbf{P})$. For $\mathbb{R}$–valued rvs X (with distribution F) and Y (with distribution G), we say that X (resp. F) is smaller than Y (resp. G) in the strong stochatic ordering (resp. convex ordering, increasing convex ordering) if

$$\mathbf{E}[\varphi(X)] \leq \mathbf{E}[\varphi(Y)] \tag{1.1}$$

for all mappings $\varphi : \mathbb{R} \to \mathbb{R}$ which are monotone increasing (resp. convex, increasing and convex) provided the expectations in (1.1) exist. In that case we write $X \leq_{st} Y$ (resp. $X \leq_{cx} Y$, $X \leq_{icx} Y$), or equivalently, $F \leq_{st} G$ (resp. $F \leq_{cx} G$, $F \leq_{icx} G$). Additional material on these orderings can be found in the monographs by Ross [10] and Stoyan [12].

We now present two characterizations of the increasing convex ordering $\leq_{icx}$ in terms of the stronger orderings $\leq_{st}$ and $\leq_{cx}$. The proof is given in Section 3, and follows in an elementary fashion from a well–known sample path characterization of $\leq_{icx}$ due to Strassen [13].

Theorem 1. *Let F and G be two integrable probability distributions on $\mathbb{R}$. The comparison $F \leq_{icx} G$ holds*

1. *if and only if there exists a probability distribution function H on $\mathbb{R}$ such that $F \leq_{cx} H$ and $H \leq_{st} G$;*

2. *if and only if there exists a probability distribution function K on $\mathbb{R}$ such that $F \leq_{st} K$ and $K \leq_{cx} G$.*

To the best of the author's knowledge, this elementary result seems not to have appeared before in the literature. In addition to showing that in some sense the ordering $\leq_{icx}$ is determined by the stronger orderings $\leq_{st}$ and $\leq_{cx}$, Theorem 1 provides a natural means for establishing the following type of comparability results: For $k = 1, 2$, let $Z^{(k)}$ denote a performance measure associated with a system characterized by the probability distribution $F^{(k)}$ in the sense that $Z^{(k)} =_{st} T(F^{(k)})$ for some transformation T. In some instances, it is expected that the comparison $F^{(1)} \leq_{icx} F^{(2)}$ implies $Z^{(1)} \leq Z^{(2)}$, with $\leq$ denoting here some appropriate notion of ordering on probability distributions. In view of Theorem 1, the inequality $Z^{(1)} \leq Z^{(2)}$ needs only be shown under each one of the *stronger* assumptions $F^{(1)} \leq_{st} F^{(2)}$ and $F^{(1)} \leq_{cx} F^{(2)}$, a fact that sometimes can be put to advantage. In Section 2, we use this two–step approach to provide a very short proof of a comparison result for $M|GI|1$ queues that Daley and Rolski [5] obtained by completely different means.

A clue as to the validity of Theorem 1 can be found in a correspondence between the stochastic ordering $\leq_{cx}$ (resp. $\leq_{icx}$) and the majorization ordering $\prec$ (resp. submajorization ordering $\prec_w$) which was first developed by Karamata [6] [8, p. 17]. As we briefly discuss it in Section 4, this allows for the possibility of mapping properties of (sub)majorization into properties of the stochastic orderings $\leq_{cx}$ and $\leq_{icx}$. In fact, Theorem 1 can be interpreted as a translation to these stochastic orderings of well- known properties of (sub)majorization [8, A.9, p. 123] which are now presented for easy comparison. Here and throughout, for elements $\mathbf{x}$ and $\mathbf{y}$ in $I\!R^k$, we write $\mathbf{x} \leq \mathbf{y}$ with the understanding that $x_i \leq y_i$, $i = 1, \ldots, k$.

Theorem 2. *For elements* $\mathbf{x}$ *and* $\mathbf{y}$ *of* $I\!R^k$, *we have* $\mathbf{x} \prec_w \mathbf{y}$

1. *if and only if there exists* $\mathbf{u}$ *in* $I\!R^k$ *such that* $\mathbf{x} \prec \mathbf{u}$ *and* $\mathbf{u} \leq \mathbf{y}$*;*
2. *if and only if there exists* $\mathbf{v}$ *in* $I\!R^k$ *such that* $\mathbf{x} \leq \mathbf{v}$ *and* $\mathbf{v} \prec \mathbf{y}$.

In [3] Chang has recently taken a similar approach in developing a new characterization of the ordering $\leq_{cx}$ through a novel cut criterion for majorization.

2. AN APPLICATION

The queueing systems considered here are $GI|GI|c$ queues with first–come first–served queueing discipline; their generic interarrival time rv T and service time rv S are square integrable. The system is said to be stable whenever $\mathbf{E}[S] < c\,\mathbf{E}[T]$, in which case an integrable stationary waiting time rv W exists [12, p. 101].

In [5], Daley and Rolski strengthened an earlier result of Rolski and Stoyan [11] concerning the comparison of $M|GI|1$ queues, and with the help of Theorem 1, we shall now revisit these results. We first summarize what was known up to the writing of [5] concerning the comparison of the stationary waiting time rv in stable $M|GI|1$ queues with respect to variations in service time distributions.

Theorem 3. *Consider two stable* $M|GI|1$ *queues with generic service time rvs* $S^{(i)}$, $i = 1, 2$, *and arrival rates* $\lambda^{(i)}$, $i = 1, 2$, *such that* $\lambda^{(1)} \leq \lambda^{(2)}$. *If either*

$$S^{(1)} \leq_{st} S^{(2)} \tag{2.1}$$

or

$$S^{(1)} \leq_{cx} S^{(2)}, \tag{2.2}$$

then the corresponding stationary waiting time rvs $W^{(i)}$, $i = 1, 2$, *satisfy*

$$W^{(1)} \leq_{st} W^{(2)}. \tag{2.3}$$

The conclusion (2.3) under (2.1) follows readily from a well–known fact concerning the monotonicity of $GI|GI|1$ queues [12, Thm. 5.2.1, p. 90]. A proof that (2.3) holds also under (2.2) appeared first in [11]. The contribution of Daley and Rolski in [5] was to show that (2.3) still holds under a condition which is strictly weaker than either (2.1) or (2.2).

Theorem 4. *The conclusion (2.3) of Theorem 2 still holds whenever*

$$S^{(1)} \leq_{icx} S^{(2)}. \tag{2.4}$$

Daley and Rolski provided two different proofs for this result; the first one is based on a probabilistic representation of the stationary waiting time rv in the $M|GI|1$ queue [7, p. 201] while the second one is analytic in flavor. Below we show how the result of Theorem 4 can be extracted very simply from Theorem 3 by invoking the characterizations of Theorem 1. To put it differently, the conclusion (2.3) under (2.4) is already contained in the older results of Theorem 3, and can be extracted from them through a basic property of the stochastic ordering $\leq_{icx}$.

A proof of Theorem 4. From Theorem 5.2.3a of [12, p. 82] there is no loss of generality in assuming $\lambda^{(1)} = \lambda^{(2)}$, as we do for the remainder of the proof. By Claim 2 of Theorem 1, we see that condition (2.4) is equivalent to the existence of an $\mathbb{R}$–valued rv $S^{(3)}$ such that

$$S^{(1)} \leq_{st} S^{(3)} \quad \text{and} \quad S^{(3)} \leq_{cx} S^{(2)}. \tag{2.5}$$

Since $S^{(1)}$ is a non–negative rv, we see from the first inequality in (2.5) that $S^{(3)}$ can be chosen non–negative, and is thus an appropriate candidate for a generic service time rv for a $M|GI|1$ system; this system is stable in view of the second inequality in (2.5). With an obvious notation, we apply Theorem 3 with (2.5) to conclude that $W^{(1)} \leq_{st} W^{(3)}$ and $W^{(3)} \leq_{st} W^{(2)}$, and the result (2.3) follows. ■

Theorem 1 also sheds some light on a conjecture made in [5] concerning stable $GI|M|c$. We begin by recalling a result obtained by Daley and Rolski [5, Prop. 4, p. 888]

Theorem 5. *For stable $GI|M|c$ queues with generic interarrival time rvs $T^{(i)}$, $i = 1, 2$, and service rates $\mu^{(i)}$, $i = 1, 2$, satisfying $\mu^{(1)} \geq \mu^{(2)}$, we have $W^{(1)} \leq_{icx} W^{(2)}$ if*

$$-T^{(1)} \leq_{icx} -T^{(2)}. \tag{2.6}$$

Daley and Rolski conjectured [5, p. 893] that the stronger result (2.3) also holds under (2.6). From well–known results on multiserver queues [12, Thm 6.2.1, p. 104] we already have (2.3) under the condition $-T^{(1)} \leq_{st} -T^{(2)}$. Therefore, in view of Claim 1 of Theorem 1, it suffices to investigate the conjecture when $-T^{(1)} \leq_{cx} -T^{(2)}$ (or equivalently, $T^{(1)} \leq_{cx} T^{(2)}$), a stronger condition which could be used to advantage during the discussion.

3. A PROOF OF THEOREM 1

We begin by a sample path characterization of the ordering $\leq_{icx}$ which derives from a celebrated result of Strassen [13].

Lemma 6. *Let F and G be two integrable probability distributions on $\mathbb{R}$. The comparison $F \leq_{icx} G$ holds if and only if there exist a probability triple, say $(\Omega, \mathcal{F}, \mathbf{P})$, and a pair of $\mathbb{R}$–valued rvs, say X and Y, defined on it and distributed according to F and G, respectively, such that*

$$X \leq \mathbf{E}[Y|X] \quad \mathbf{P} - a.s. \tag{3.1}$$

A proof of Theorem 1. The sufficiency part in both Claims of Theorem 1 is clear since the stochastic orders $\leq_{cx}$ and $\leq_{st}$ are each stronger than $\leq_{icx}$. Therefore, it remains only to show the necessity part of the results. To do so. we invoke the representation of Lemma 6, and with the notation introduced there, we define the rvs U and V by

$$U := X - (\mathbf{E}[Y|X] - Y) \quad \text{and} \quad V := \mathbf{E}[Y|X]. \tag{3.2}$$

(Claim 1) From (3.1)–(3.2) we readily see that

$$U \leq Y \quad \text{and} \quad \mathbf{E}[U|X] = X \quad \mathbf{P} - a.s. \tag{3.3}$$

Therefore, we have $U \leq_{st} Y$ by coupling [12, Prop. 1.2.1, p. 4] and $X \leq_{cx} U$ by Jensen's inequality, whence the desired distribution H can be taken to be that of U.

(Claim 2) First we find $X \leq_{st} V$ by making use of (3.1). Next, the integrable rv V being $\sigma(X)$–measurable, we can invoke the smoothing property of conditional expectations to get

$$\mathbf{E}[Y|V] = \mathbf{E}[\mathbf{E}[Y|X]|V] = \mathbf{E}[V|V] = V \quad \mathbf{P} - a.s. \tag{3.4}$$

Therefore, as before, we obtain $V \leq_{cx} Y$ by Jensen's inequality and the desired distribution K can be taken to be that of V. ∎

4. KARAMATA'S DEVICE

Let $\mathcal{D}$ denote the class of all distributions on $I\!R$ with *finite* support and *rational* discontinuities. A probability distribution F in $\mathcal{D}$ is characterized by a finite set of distinct points $\xi_1, \ldots, \xi_p$ on the real line $I\!R$, and by positive rational numbers $f_1, \ldots, f_p$ such that for any rv X with distribution F, we have

$$\mathbf{P}[X = \xi_i] = f_i, \quad i = 1, \ldots, p. \tag{4.1}$$

Obviously, since $0 < f_i < 1$, $i = 1, \ldots, p$, and $f_1 + \ldots + f_p = 1$, the rational probabilities $f_1, \ldots, f_p$ can be represented in the form $f_i = \frac{k_i}{k}$, $i = 1, \ldots, p$, with positive integers $k_1, \ldots, k_p, k$ related by $k = k_1 + \ldots + k_p$. Such an integer representation is of course not unique since it is determined only up to an integer factor, i.e., $f_i = \frac{ck_i}{ck}$, $i = 1, \ldots, p$, for any positive integer c. Now, with any such integer representation of the distribution F we can associate an element $\mathbf{x} = (x_1, \ldots, x_k)$ of $I\!R^k$ by repeating ξ_i as a component exactly k_i times, $i = 1, \ldots, p$, i.e.,

$$\mathbf{x} := (\underbrace{\xi_1, \ldots, \xi_1}_{k_1}, \ldots, \underbrace{\xi_i, \ldots, \xi_i}_{k_i}, \ldots, \underbrace{\xi_p, \ldots, \xi_p}_{k_p}). \tag{4.2}$$

We call this vector $\mathbf{x}$ the Karamata representation (of dimension k) of the distribution F (in $\mathcal{D}$). Several obvious but useful comments are in order concerning this representation:

(R1) With X still denoting any $I\!R$–valued rv with distribution F, we can write

$$\mathbf{E}[\varphi(X)] = \frac{1}{k} \sum_{i=1}^{k} \varphi(x_i) \tag{4.3}$$

for any mapping $\varphi : I\!R \to I\!R$;

(R2) Any element $\mathbf{x}$ of $I\!R^k$ will define a distribution F in $\mathcal{D}$. To see that, let $\xi_1, \ldots, \xi_p$ denote the distinct elements among the components $x_1, \ldots, x_k$ of $\mathbf{x}$, and let k_i denote the number of times that ξ_i occurs among these components, $i = 1, \ldots, p$. The rational numbers $f_i = \frac{k_i}{k}$, $i = 1, \ldots, p$, constitute a probability mass vector which can be used to define a probability distribution with support on the finite set of points $\xi_1, \ldots, \xi_p$ and with rational discontinuities at these points.

As should be plain from the proof of the following special case of Theorem 1, Theorem 2 is subsumed by Theorem 1, and is indeed suggestive of it.

Theorem 1bis. *Theorem 1 holds whenever F and G both belong to $\mathcal{D}$.*

Proof. In the interest of brevity we only consider Claim 2, as the discussion of Claim 1 follows the same pattern. Let F and G be two distributions in $\mathcal{D}$ with Karamata representations $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^k$, respectively. That the Karamata representations associated with F and G can always be chosen to be of the same dimension follows from the fact that the integer representation of an element of $\mathcal{D}$ is unique only up to a multiplicative constant.

From **(R1)**, the condition $F \leq_{icx} G$ is seen via (1.1) to be equivalent to

$$\sum_{i=1}^{k} \varphi(x_i) \leq \sum_{i=1}^{k} \varphi(y_i) \tag{4.4}$$

for *all* increasing convex mappings $\varphi : \mathbb{R} \to \mathbb{R}$. But this statement is known [8, B.2, p. 109] to be a necessary and sufficient condition for $\mathbf{x} \prec_w \mathbf{y}$ (on $\mathbb{R}^k$). On the other hand, by Claim **2** of Theorem 2, there exists a vector $\mathbf{v}$ of $\mathbb{R}^k$ such that $\mathbf{x} \leq \mathbf{v}$ and $\mathbf{v} \prec \mathbf{y}$. As pointed out earlier in **(R2)**, $\mathbf{v}$ is the Karamata representation of a distribution K in $\mathcal{D}$, which we now show is the requisite distribution. For future use, let X, Y and V denote rvs distributed according to F, G and K, respectively.

Since $\mathbf{x} \leq \mathbf{v}$, we have for any increasing mapping $\varphi : \mathbb{R} \to \mathbb{R}$ that $\varphi(x_i) \leq \varphi(v_i)$, $i = 1, \ldots, k$, whence $\mathbf{E}[\varphi(X)] \leq \mathbf{E}[\varphi(V)]$ by **(R1)** and the comparison $X \leq_{st} V$ is obtained. Next, by a well-known result of Hardy, Littlewood and Pólya [8, B.1., p. 108], we have that $\mathbf{v} \prec \mathbf{y}$ is equivalent to

$$\sum_{i=1}^{k} \varphi(v_i) \leq \sum_{i=1}^{k} \varphi(y_i) \tag{4.5}$$

for *all* convex mappings $\varphi : \mathbb{R} \to \mathbb{R}$. The conclusion $V \leq_{cx} Y$ is now an immediate consequence of the remark **(R1)** and of the definition of $\leq_{cx}$. ∎

Theorem 1bis points in principle to an alternate proof of Theorem 1, whereby the result first obtained for distributions in $\mathcal{D}$, is extended to the class of all probability distributions on $\mathbb{R}$ by a multi-step approximation procedure via limiting arguments.

5. ACKNOWLEDGEMENTS

The work reported here was started during a short stay at INRIA, Sophia Antipolis. The hospitality of Project MEVAL is gratefully acknowledged. In particular the author wishes to thank Zhen Liu for several discussions on the topic of this paper.

REFERENCES

[1] F. Baccelli and A. M. Makowski, "Queueing models for systems with synchronization constraints," *Proceedings of the IEEE* **77** (1989), Invited paper, Special Issue on Dynamics of Discrete Event Systems, pp. 138–161.

[3] C.–S. Chang, "Smoothing point processes as a means to increase throughput," IBM Research Report **RC–16866** (1991)

[5] D.J. Daley and T. Rolski, "Some comparison results for waiting times in single– and many–server queues," *Journal of Applied Probability* **21** (1984), pp. 887–900.

[6] J. Karamata, "Sur une inégalité relative aux fonctions convexes," *Publications Mathématiques de l'Université de Belgrade* **1** (1932), pp. 145–148.

[7] L. Kleinrock, *Queueing Systems I: Theory*, J. Wiley & Sons, New York (NY) (1976).

[8] A.W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, New York (NY) (1979).

[11] T. Rolski and D. Stoyan, "On the comparison of waiting times in $GI|G|1$ queues," *Operations Resaerch* **24** (1976), pp. 197–200.

[10] S. Ross, *Stochastic Processes*, J. Wiley & Sons, New York (NY) (1984).

[12] D. Stoyan, *Comparison Methods for Queues and Other Stochastic Models*, English Translation (D.J. Daley, Editor), J. Wiley & Sons, New York (NY) (1984).

[13] V. Strassen, "The existence of probability measures with given marginals," *Annals of Matematical Statistics* **36** (1965), pp. 423–439.