Electronic Journal of Probability

Vol. 14 (2009), Paper no. 26, pages 663–727.

Journal URL
http://www.math.washington.edu/~ejpecp/

Optimal two-value zero-mean disintegration of zero-mean random variables

Iosif Pinelis*
Department of Mathematical Sciences
Michigan Technological University
Houghton, Michigan 49931
email ipinelis@mtu.edu

Abstract

For any continuous zero-mean random variable (r.v.) X, a *reciprocating* function r is constructed, based only on the distribution of X, such that the conditional distribution of X given the (at-most-)two-point set $\{X, \mathsf{r}(X)\}$ is the zero-mean distribution on this set; in fact, a more general construction without the continuity assumption is given in this paper, as well as a large variety of other related results, including characterizations of the reciprocating function and modeling distribution asymmetry patterns. The mentioned disintegration of zero-mean r.v.'s implies, in particular, that an arbitrary zero-mean distribution is represented as the mixture of two-point zero-mean distributions; moreover, this mixture representation is most symmetric in a variety of senses.

Somewhat similar representations – of any probability distribution as the mixture of two-point distributions with the same skewness coefficient (but possibly with different means) – go back to Kolmogorov; very recently, Aizenman *et al.* further developed such representations and applied them to (anti-)concentration inequalities for functions of independent random variables and to spectral localization for random Schroedinger operators. One kind of application given in the present paper is to construct certain statistical tests for asymmetry patterns and for location without symmetry conditions. Exact inequalities implying conservative properties of such tests are presented. These developments extend results established earlier by Efron, Eaton, and Pinelis under a symmetry condition.

*Supported in part by NSF grant DMS-0805946

Key words: Disintegration of measures, Wasserstein metric, Kantorovich-Rubinstein theorem, transportation of measures, optimal matching, most symmetric, hypothesis testing, confidence regions, Student's t-test, asymmetry, exact inequalities, conservative properties.

AMS 2000 Subject Classification: Primary 28A50, 60E05, 60E15, 62G10, 62G15, 62F03, 62F25; Secondary: 49K30, 49K45, 49N15, 60G50, 62G35, 62G09, 90C08, 90C46.

Submitted to EJP on March 27, 2008, final version accepted March 4, 2009.

1 Introduction

Efron [11] observed that the tail of the distribution of the self-normalized sum

$$S := \frac{X_1 + \cdots + X_n}{\sqrt{X_1^2 + \cdots + X_n^2}}, \tag{1.1}$$

is bounded in certain sense by the tail of the standard normal distribution — provided that the X_i's satisfy a certain symmetry condition; it is enough that the X_i's be independent and symmetrically (but not necessarily identically) distributed. Thus, one obtains a conservative test for symmetry. Further results under such a symmetry condition were obtained by Eaton [8; 9], Eaton and Efron [10], and Pinelis [24]. Note that the self-normalized sum S is equivalent to the t statistic, in the sense that they can be obtained from each other by a monotonic transformation.

A simple but crucial observation made by Efron in [11] was that the conditional distribution of any symmetric r.v. X given $|X|$ is the symmetric distribution on the (at-most-)two-point set $\{|X|, -|X|\}$. Therefore, under the symmetry condition, the distribution of the self-normalized sum is a mixture of the distributions of the normalized Khinchin-Rademacher sums, and thus can be nicely bounded, say by using an exponential bound by Hoeffding [17], or more precise bounds by Eaton [8; 9] and Pinelis [24].

However, the mentioned results do not hold in general without a symmetry assumption. In fact, as pointed out already by Bartlett [4] and confirmed by Ratcliffe [30], asymmetry may quite significantly affect the t distribution, in a sense more than kurtosis may. These results are in agreement with the one by Hall and Wang [14]. Tukey [34, page 206] wrote, "It would be highly desirable to have a modified version of the t-test with a greater resistance to skewness... ." This concern will be addressed in the present paper by such results as Corollaries 2.5 and 2.6.

Closely related to this is the question of modeling asymmetry. Tukey [35] proposed using the power-like transformation functions of the form $z(y) = a(y+c)^p + b$, $y > -c$, with the purpose of symmetrizing the data. To deal with asymmetry and heavy tails, Tukey also proposed (see Kafadar [19, page 328] and Hoaglin [15]) the so-called g-h technology, whereby to fit the data to a g-h distribution, which is the distribution of a r.v. of the form $e^{hZ^2/2}(e^{gZ}-1)/g$, where $Z \sim N(0,1)$, so that the parameters g and h are responsible, respectively, for the skewness of the distribution and the heaviness of the tails. In this paper, we propose modeling asymmetry using what we shall refer to as *reciprocating functions*.

The basic idea here is to represent any zero-mean, possibly asymmetric, distribution as an appropriate mixture of two-point zero-mean distributions. To present this idea quickly, let us assume at this point that a zero-mean r.v. X has an everywhere continuous and strictly increasing distribution function (d.f.). Consider the truncated r.v. $\tilde{X}_{a,b} := X\,\mathrm{I}\{a \leqslant X \leqslant b\}$. (Here and in what follows $\mathrm{I}\{A\}$ stands, as usual, for the indicator of a given assertion A, so that $\mathrm{I}\{A\} = 1$ if A is true and $\mathrm{I}\{A\} = 0$ if A is false.) Then, for every fixed $a \in (-\infty, 0]$, the function $b \mapsto \mathsf{E}\tilde{X}_{a,b}$ is continuous and increasing on the interval $[0,\infty)$ from $\mathsf{E}\tilde{X}_{a,0} \leqslant 0$ to $\mathsf{E}\tilde{X}_{a,\infty} > 0$. Hence, for each $a \in (-\infty, 0]$, there exists a unique value $b \in [0,\infty)$ such that $\mathsf{E}\tilde{X}_{a,b} = 0$. Similarly, for each $b \in [0,\infty)$, there exists a unique value $a \in (-\infty, 0]$ such that $\mathsf{E}\tilde{X}_{a,b} = 0$. That is, one has a one-to-one correspondence between $a \in (-\infty, 0]$ and $b \in [0,\infty)$ such that $\mathsf{E}\tilde{X}_{a,b} = 0$. Denote by $\mathsf{r} = \mathsf{r}_X$ the *reciprocating* function defined on $\mathbb{R}$ and carrying this correspondence, so that

$$\mathsf{E}X\,\mathrm{I}\{X \text{ is between } x \text{ and } \mathsf{r}(x)\} = 0 \quad \forall x \in \mathbb{R};$$

the function r is decreasing on $\mathbb{R}$ and such that $\mathsf{r}(\mathsf{r}(x)) = x\ \forall x \in \mathbb{R}$; moreover, $\mathsf{r}(0) = 0$. (Clearly, $\mathsf{r}(x) = -x$ for all real x if the r.v. X is also symmetric.) Thus, the set $\{\, \{x, \mathsf{r}(x)\} \colon x \in \mathbb{R}\, \}$ of two-point sets constitutes a partition of $\mathbb{R}$. One can see that the conditional distribution of the zero-mean r.v. X given the random two-point set $\{X, \mathsf{r}(X)\}$ is the uniquely determined zero-mean distribution on the set $\{X, \mathsf{r}(X)\}$.

It follows that the distribution of the zero-mean r.v. X with a continuous strictly increasing d.f. is represented as a mixture of two-point zero-mean distributions. A somewhat similar representation — of any probability distribution as the mixture of two-point distributions with the same skewness coefficient $\frac{q-p}{\sqrt{pq}}$ (but possibly with different means) – goes back to Kolmogorov; very recently Aizenman *et al.* [3] further developed this representation and applied it to (anti-)concentration inequalities for functions of independent random variables and to spectral localization for random Schroedinger operators.

In accordance with their purposes, instead of r.v.'s $\tilde{X}_{a,b} = X\,\mathrm{I}\{a \leqslant X \leqslant b\}$ Aizenman *et al.* [3] (who refer to a and b as markers) essentially deal with r.v.'s (i) $\mathrm{I}\{X \leqslant a\} - \mathrm{I}\{X > b\}$ (in a case of markers moving in opposite directions) and with (ii) $\mathrm{I}\{X \leqslant a\} - \mathrm{I}\{q_{1-p} < X \leqslant b\}$ (in a case of markers moving in the same direction, where q_{1-p} is a $(1-p)$-quantile of the distribution of X). The construction described above in terms of $\tilde{X}_{a,b} = X\,\mathrm{I}\{a \leqslant X \leqslant b\}$ corresponds, clearly, to the case of opposite-moving markers.

While an analogous same-direction zero-mean disintegration is possible, we shall not deal with it in this paper. For a zero-mean distribution, the advantage of an opposite-directions construction is that the resulting two-point zero-mean distributions are less asymmetric than those obtained by using a same-direction method (in fact, we shall show that our opposite-directions disintegration is most symmetric, in a variety of senses). On the other hand, the same-direction method will produce two-point zero-mean distributions that are more similar to one another in width. Thus, in our main applications – to self-normalized sums, the advantages of opposite-directions appear to be more important, since the distribution of a self-normalized sum is much more sensitive to the asymmetry than to the inhomogeneity of the constituent two-point distributions in width; this appears to matter more in the setting of Corollary 2.6 than in the one of Corollary 2.5.

These mixture representations of a distribution are similar to the representations of the points of a convex compact set as mixtures of the extreme points of the set; the existence of such representations is provided by the celebrated Krein-Milman-Choquet-Bishop-de Leeuw (KMCBdL) theorem; concerning "non-compact" versions of this theorem see e.g. [22]. In our case, the convex set would be the set of all zero-mean distributions on $\mathbb{R}$. However, in contrast with the KMCBdL-type pure-existence theorems, the representations given in [27], [3], and this paper are constructive, specific, and, as shown here, optimal, in a variety of senses.

Moreover, in a certain sense [27] and this paper provide disintegration of r.v.'s rather than that of their distributions, as the two-point set $\{x, \mathsf{r}(x)\}$ is a function of the observed value x of the r.v. X. This makes it convenient to construct statistical tests for asymmetry patterns and for location without symmetry conditions. Exact inequalities implying conservative properties of such tests will be given in this paper. These developments extend the mentioned results established earlier by Efron, Eaton, and Pinelis under the symmetry condition.

More specifically, one can construct generalized versions of the self-normalized sum (1.1), which require – instead of the symmetry of independent r.v.'s X_i – only that the X_i's be zero-mean:

$$S_W := \frac{X_1 + \cdots + X_n}{\frac{1}{2}\sqrt{W_1^2 + \cdots + W_n^2}} \quad \text{and} \quad S_{Y,\lambda} := \frac{X_1 + \cdots + X_n}{(Y_1^\lambda + \cdots + Y_n^\lambda)^{\frac{1}{2\lambda}}},$$

where $\lambda > 0$, $W_i := |X_i - \mathsf{r}_i(X_i)|$ and $Y_i := |X_i\, \mathsf{r}_i(X_i)|$, and the reciprocating function $\mathsf{r}_i := \mathsf{r}_{X_i}$ is constructed as above, based on the distribution of X_i, for each i, so that the r_i's may be different from one another if the X_i's are not identically distributed. Note that $S_W = S_{Y,1} = S$ (recall here (1.1)) when the X_i's are symmetric. Logan *et al* [21] and Shao [31] obtained limit theorems for the "symmetric" version of $S_{Y,\lambda}$ (with X_i^2 in place of Y_i), whereas the X_i's were not assumed to be symmetric.

Corollaries 2.5 and 2.6 in Subsection 2.2 of this paper suggest that statistical tests based on the "corrected for asymmetry" statistics S_W and S_Y have desirable conservativeness and similarity properties, which could result in greater power; further studies are needed here. (Recall that a test is referred to as (approximately) *similar* if the type I error probabilities are (approximately) the same for all distributions corresponding to the null hypothesis.)

Actually, in this paper we provide two-point zero-mean disintegration of any zero-mean r.v. X, with a d.f. not necessarily continuous or strictly increasing. Toward that end, randomization (by means of a r.v. uniformly distributed in interval $(0,1)$) is used to deal with the atoms of the distribution of r.v. X, and generalized inverse functions to deal with the intervals on which the d.f. of X is constant.

Note that the reciprocating function r depends on the usually unknown in statistics distribution of the underlying r.v. X. However, if e.g. the X_i's constitute an i.i.d. sample, then the function G defined in the next section by (2.1) can be estimated based on the sample, so that one can estimate the reciprocating function r. Thus, replacing $X_1 + \cdots + X_n$ in the numerators of S_W and $S_{Y,\lambda}$ by $X_1 + \cdots + X_n - n\theta$, one obtains approximate pivots to be used to construct confidence intervals or, equivalently, tests for an unknown mean θ. One can use bootstrap to estimate the distributions of such approximate pivots. The asymmetry-adjusted self-normalized sums can be also used to *test for asymmetry patterns*, which, as stated, may be conveniently and flexibly modeled by reciprocating functions.

In testing for the mean of the unknown distribution, the asymmetry pattern as represented by the reciprocating function should be considered as a nuisance parameter. The modifications S_W and S_Y of the self-normalized sum serve the purpose of removing or minimizing the dependence of the distribution of the test statistic on the asymmetry pattern. Such modifications are quite common in statistics. For instance, the very idea of "Studentization" or self-normalization is of this kind, since such a procedure removes or minimizes the dependence of the distribution of the test statistic on the variance, which latter is a nuisance parameter when testing for the mean.

As was noted, the reciprocating functions in the expressions for S_W and S_Y are usually to be estimated based on the sample. Such a step — replacing nuisance parameters by their empirical counterparts — is also quite common in statistics. A standard example is testing for the variance σ^2 in the simple regression model $Y_i = \alpha + \beta x_i + \varepsilon_i$ $(i = 1, \dots, n)$, where $x_1, \dots, x_n$ are given real numbers, $\varepsilon_1, \dots, \varepsilon_n$ are independent identically distributed (i.i.d.) r.v.'s with mean 0 and variance σ^2, and α, β, σ^2 are unknown parameters. Then the construction of a test statistic for σ^2 can be considered as a *two-step procedure*. First, the actually available observations Y_i are replaced by $Y_i - \alpha - \beta x_i$, which are of course the ε_i's, whose joint distribution does not depend on the nuisance parameters α and β. However, since α and β are unknown, a second step is needed, for α and β to be replaced by some estimates, say $\hat\alpha$ and $\hat\beta$, based on the observations Y_i. This two-step procedure results in the residuals $R_i := Y_i - \hat\alpha - \hat\beta x_i$, which one can hope to be (at least for large n) approximately i.i.d. with approximate mean 0 and approximate variance σ^2. Then the distribution of the statistic $\hat S^2$ defined as $\frac{1}{n}\sum_{i=1}^n (R_i - \overline{R})^2$ or $\frac{1}{n-2}\sum_{i=1}^n (R_i - \overline{R})^2$ (with $\overline{R} := \frac{1}{n}\sum_{i=1}^n R_i$) will not depend on the nuisance parameters α and β if the ε_i are normally distributed, and otherwise will usually depend only slightly on α and β, again if n is large. Thus, $\hat S^2$ can be used as an estimator of σ^2 or the corresponding test statistic.

Going back to the problem of adjusting the t statistic for skewness, the procedure used e.g. by Hall [13, page 123] follows the same general two-step approach as the one outlined above; the first step of this procedure is based on an Edgeworth expansion; and the second step, on a corresponding empirical estimate of a skewness parameter. The mentioned g-h technology by Tukey — first to try to remove the skewness and kurtosis by applying a g-h transformation to the sample and then to estimate the nuisance parameters g and h in order to find a best fit for the data — is also of the same general two-step kind.

With our method, the nuisance parameter as represented by the reciprocating function may be finite- or infinite-dimensional, corresponding to a parametric, semi-parametric, or nonparametric model — see Subsection 3.5 below on modeling. In contrast, the methods used by Hall and Tukey allow only for a one-dimensional nuisance parameter for skewness. Thus, modeling with reciprocating functions appears to provide more flexibility. It also allows the mentioned conservative properties of the tests to be preserved without a symmetry condition.

Quite different approaches to the problem of asymmetry were demonstrated in such papers as [5; 26; 28; 6], where the authors assessed, under certain conditions, the largest possible effect that asymmetry may cause.

This paper is an improvement of preprint [27]: the results are now much more numerous and comprehensive, and also somewhat more general, while the proof of the basic result (done here using a completely different method) is significantly shorter. A brief account of results of [27] (without proofs) was presented in [28].

2 Statements of main results on disintegration

2.1 Two-value zero-mean disintegration of one zero-mean r.v.

Let ν be any (nonnegative finite) measure defined on $\mathscr{B}(\mathbb{R})$, where $\mathscr{B}(E)$ stands for the set of all Borel subsets of a given set E. Sometimes it will be convenient to consider such a measure ν extended to $\mathscr{B}([-\infty,\infty])$ so that, naturally, $\nu(\{-\infty\}) = \nu(\{\infty\}) = 0$. Consider the function $G = G_\nu$ with values in $[0,\infty]$ defined by the formula

$$G(x) := G_\nu(x) := \begin{cases} \int_{(0,x]} z\,\nu(\mathrm{d}z) & \text{if } x \in [0,\infty], \\ \int_{[x,0)} (-z)\,\nu(\mathrm{d}z) & \text{if } x \in [-\infty,0]. \end{cases} \tag{2.1}$$

Note that

$$\begin{gathered} G(0) = 0;\ G \text{ is non-decreasing on } [0,\infty] \text{ and right-continuous on } [0,\infty); \\ \text{and } G \text{ is non-increasing on } [-\infty,0] \text{ and left-continuous on } (-\infty,0]; \end{gathered} \tag{2.2}$$

in particular, G is continuous at 0.

Define next the positive and negative generalized inverses x_+ and x_- of the function G:

$$x_+(h) := x_{+,\nu}(h) := \inf\{x \in [0,\infty] \colon G_\nu(x) \geqslant h\}, \tag{2.3}$$
$$x_-(h) := x_{-,\nu}(h) := \sup\{x \in [-\infty,0] \colon G_\nu(x) \geqslant h\}, \tag{2.4}$$

for any $h \in [-\infty,\infty]$; here, as usual, $\inf\emptyset := \infty$ and $\sup\emptyset := -\infty$.

Introduce also a "randomized" version of G:

$$\tilde{G}(x,u) := \tilde{G}_\nu(x,u) := \begin{cases} G_\nu(x-) + (G_\nu(x) - G_\nu(x-))\,u & \text{if } x \in [0,\infty], \\ G_\nu(x+) + (G_\nu(x) - G_\nu(x+))\,u & \text{if } x \in [-\infty,0] \end{cases} \tag{2.5}$$

and what we shall refer to as the *reciprocating* function $\mathsf{r} = \mathsf{r}_\nu$ for the measure ν:

$$\mathsf{r}(x,u) := \mathsf{r}_\nu(x,u) := \begin{cases} x_{-,\nu}(\tilde{G}_\nu(x,u)) & \text{if } x \in [0,\infty], \\ x_{+,\nu}(\tilde{G}_\nu(x,u)) & \text{if } x \in [-\infty,0], \end{cases} \tag{2.6}$$

for all $u \in [0,1]$.

In the case when ν is a non-atomic probability distribution, a similar construction of the reciprocating function (without using this term) was given already by Skorokhod in the proof of [32, Lemma 3]; cf. [20, Lemma 14.4]; thanks are due to Lutz Mattner for the latter reference and to Olav Kallenberg for the reference to Skorohod.

Remark 2.1. (i) The function $\tilde{G}$ is Borel(-measurable), since each of the functions G, $G(\cdot+)$, $G(\cdot-)$ is monotonic on $[0,\infty)$ and $(-\infty,0]$ and hence Borel. Therefore and by property (i) of Proposition 3.1, stated in the next section, the reciprocating function r is Borel, too.

(ii) Also, $\tilde{G}(x,u)$ and hence $\mathsf{r}(x,u)$ depend on u for a given value of x only if $\nu(\{x\}) \neq 0$. Therefore, let us write simply $\mathsf{r}(x)$ in place of $\mathsf{r}(x,u)$ in the case when the measure ν is non-atomic.

If ν is the measure $\mu = \mu_X$ that is the distribution of a r.v. X, then we may use subscript $_X$ with G, $\tilde{G}$, r, $x_\pm$ in place of subscript $_\mu$ (or no subscript at all).

In what follows, X will by default denote an arbitrary *zero-mean* real-valued r.v., which will be usually thought of as fixed. Then, for $G = G_X$,

$$G(\infty) = G(-\infty) = G(\infty-) = G\big((-\infty)+\big) = \tfrac{1}{2}\,\mathsf{E}\,|X| =: m < \infty. \tag{2.7}$$

Let U stand for any r.v. which is independent of X and uniformly distributed on the unit interval $[0,1]$.

For any a and b in $\mathbb{R}$ such that $ab \leqslant 0$, let $X_{a,b}$ denote any *zero-mean* r.v. with values in the two-point set $\{a,b\}$; note that such a r.v. $X_{a,b}$ exists and, moreover, its distribution is uniquely determined:

$$\mathsf{P}(X_{a,b} = a) = \tfrac{b}{b-a} \quad \text{and} \quad \mathsf{P}(X_{a,b} = b) = \tfrac{a}{a-b} \tag{2.8}$$

if $a \neq b$, and $X_{a,b} = 0$ almost surely (a.s.) if $a = b (= 0)$; then in fact $X_{a,b} = 0$ a.s. whenever $ab = 0$. Along with the r.v. $X_{a,b}$, consider

$$R_{a,b} := \mathsf{r}_{a,b}(X_{a,b}, U) \tag{2.9}$$

provided that U does not depend on $X_{a,b}$, where $\mathsf{r}_{a,b} := \mathsf{r}_{X_{a,b}}$, the reciprocal function for $X_{a,b}$. Note that, if $ab = 0$, then $R_{a,b} = 0 = X_{a,b}$ a.s. If $ab < 0$, then $R_{a,b} = b$ a.s. on the event $\{X_{a,b} = a\}$, and $R_{a,b} = a$ a.s. on the event $\{X_{a,b} = b\}$, so that the random set $\{X_{a,b}, R_{a,b}\}$ coincides a.s. with the nonrandom set $\{a, b\}$. However, $R_{a,b}$ equals in distribution to $X_{a,b}$ only if $a + b = 0$, that is, only if $X_{a,b}$ is symmetric; moreover, in contrast with $X_{a,b}$, the r.v. $R_{a,b}$ is zero-mean only if $a + b = 0$. Clearly, $(X_{a,b}, R_{a,b}) \overset{\mathrm{D}}{=} (X_{b,a}, R_{b,a})$ whenever $ab \leqslant 0$.

We shall prove that the conditional distribution of X given the two-point random set $\{X, \mathsf{r}(X,U)\}$ is the zero-mean distribution on this set:

$$\big(X \,\big|\, \{X, \mathsf{r}(X,U)\} = \{a,b\}\big) \overset{\mathrm{D}}{=} X_{a,b}. \tag{2.10}$$

In fact, we shall prove a more general result: that the conditional distribution of the ordered pair $\big(X, \mathsf{r}(X,U)\big)$ given that $\{X, \mathsf{r}(X,U)\} = \{a,b\}$ is the distribution of the ordered pair $\big(X_{a,b}, R_{a,b}\big)$:

$$\Big(\big(X, \mathsf{r}(X,U)\big) \,\Big|\, \{X, \mathsf{r}(X,U)\} = \{a,b\}\Big) \overset{\mathrm{D}}{=} \big(X_{a,b}, R_{a,b}\big). \tag{2.11}$$

Formally, this basic result of the paper is expressed as

Theorem 2.2. *Let $g\colon \mathbb{R}^2 \to \mathbb{R}$ be any Borel function bounded from below (or from above). Then*

$$\mathsf{E}\, g\big(X, \mathsf{r}(X,U)\big) = \int_{\mathbb{R}\times[0,1]} \mathsf{E}\, g\big(X_{x,\mathsf{r}(x,u)}, R_{x,\mathsf{r}(x,u)}\big)\; \mathsf{P}(X \in \mathrm{d}x)\,\mathrm{d}u. \tag{2.12}$$

Instead of the condition that g be bounded from below or above, it is enough to require only that $g(x,r) - cx$ be so for some real constant c over all real x, r.

The proofs (whenever necessary) are deferred to Section 4.

As one can see, Theorem 2.2 provides a complete description of the distribution of the ordered random pair $\big(X, \mathsf{r}(X,U)\big)$ – as a mixture of two-point distributions on $\mathbb{R}^2$; each of these two-point distributions is supported by a two-point subset of $\mathbb{R}^2$ of the form $\{(a,b),(b,a)\}$ with $ab \leqslant 0$, and at that the mean of the projection of this two-point distribution onto the first coordinate axis is zero. As special cases, Theorem 2.2 contains descriptions of the individual distributions of the r.v.'s X and $\mathsf{r}(X,U)$ as mixtures of two-point distributions on $\mathbb{R}$: for any Borel function $g\colon \mathbb{R} \to \mathbb{R}$ bounded from below (or from above) one has

$$\mathsf{E}\, g(X) = \int_{\mathbb{R}\times[0,1]} \mathsf{E}\, g\big(X_{x,\mathsf{r}(x,u)}\big)\; \mathsf{P}(X \in \mathrm{d}x)\,\mathrm{d}u; \tag{2.13}$$

$$\mathsf{E}\, g\big(\mathsf{r}(X,U)\big) = \int_{\mathbb{R}\times[0,1]} \mathsf{E}\, g\big(R_{x,\mathsf{r}(x,u)}\big)\; \mathsf{P}(X \in \mathrm{d}x)\,\mathrm{d}u.$$

This is illustrated by

Example 2.3. Let X have the discrete distribution $\frac{5}{10}\,\delta_{-1} + \frac{1}{10}\,\delta_0 + \frac{3}{10}\,\delta_1 + \frac{1}{10}\,\delta_2$ on the finite set $\{-1,0,1,2\}$, where δ_a denotes the (Dirac) probability distribution on the singleton set $\{a\}$. Then $m = \frac{5}{10}$ and, for $x \in \mathbb{R}$, $u \in [0,1]$, and $h \in [0,m]$,

$$G(x) = \tfrac{5}{10}\,\mathrm{I}\{x \leqslant -1\} + \tfrac{3}{10}\,\mathrm{I}\{1 \leqslant x < 2\} + \tfrac{5}{10}\,\mathrm{I}\{2 \leqslant x\},$$

$$x_+(h) = \mathrm{I}\{0 < h \leqslant \tfrac{3}{10}\} + 2\,\mathrm{I}\{\tfrac{3}{10} < h\}, \quad x_-(h) = -\,\mathrm{I}\{0 < h\},$$

$$\tilde{G}(-1,u) = \tfrac{5}{10}\,u, \quad \tilde{G}(0,u) = 0, \quad \tilde{G}(1,u) = \tfrac{3}{10}\,u, \quad \tilde{G}(2,u) = \tfrac{3}{10} + \tfrac{2}{10}\,u,$$

$$\mathsf{r}(-1,u) = \mathrm{I}\{u \leqslant \tfrac{3}{5}\} + 2\,\mathrm{I}\{u > \tfrac{3}{5}\}, \; \mathsf{r}(0,u) = 0, \; \mathsf{r}(1,u) = -1, \; \mathsf{r}(2,u) = -1.$$

Therefore, the distribution of the random set $\{X, \mathsf{r}(X,U)\}$ is $\frac{6}{10}\,\delta_{\{-1,1\}} + \frac{3}{10}\,\delta_{\{-1,2\}} + \frac{1}{10}\,\delta_{\{0\}}$, and the conditional distributions of X given $\{X, \mathsf{r}(X,U)\} = \{-1,1\}$, $\{X, \mathsf{r}(X,U)\} = \{-1,2\}$, and $\{X, \mathsf{r}(X,U)\} = \{0\}$ are the zero-mean distributions $\frac{1}{2}\,\delta_{-1} + \frac{1}{2}\,\delta_1$, $\frac{2}{3}\,\delta_{-1} + \frac{1}{3}\,\delta_2$, and δ_0, respectively. Thus, the zero-mean distribution of X is represented as a mixture of these two-point zero-mean distributions:

$$\tfrac{5}{10}\,\delta_{-1} + \tfrac{1}{10}\,\delta_0 + \tfrac{3}{10}\,\delta_1 + \tfrac{1}{10}\,\delta_2 = \tfrac{6}{10}\,(\tfrac{1}{2}\,\delta_{-1} + \tfrac{1}{2}\,\delta_1) + \tfrac{3}{10}\,(\tfrac{2}{3}\,\delta_{-1} + \tfrac{1}{3}\,\delta_2) + \tfrac{1}{10}\,\delta_0.$$

2.2 Two-value zero-mean disintegration of several independent zero-mean r.v.'s and applications to self-normalized sums

Suppose here that $X_1,\dots,X_n$ are independent zero-mean r.v.'s and $U_1,\dots,U_n$ are independent r.v.'s uniformly distributed on $[0,1]$, which are also independent of $X_1,\dots,X_n$. For each $j = 1,\dots,n$, let $R_j := \mathsf{r}_j(X_j, U_j)$, where r_j denotes the reciprocating function for r.v. X_j. For any real $a_1, b_1, \dots, a_n, b_n$ such that $a_j b_j \leqslant 0$ for all j, let

$$X_{1;a_1,b_1},\dots,X_{n;a_n,b_n}$$

be independent r.v.'s such that, for each $j \in \{1,\dots,n\}$, the r.v. $X_{j;a_j,b_j}$ is zero-mean and takes on its values in the two-point set $\{a_j, b_j\}$. For all j, let

$$R_{j;a_j,b_j} := a_j\, b_j / X_{j;a_j,b_j}$$

if $a_j b_j < 0$ and $R_{j;a_j,b_j} := 0$ if $a_j b_j = 0$.

Theorem 2.4. *Let $g\colon \mathbb{R}^{2n} \to \mathbb{R}$ be any Borel function bounded from below (or from above). Then identity* (2.12) *can be generalized as follows:*

$$\mathsf{E}\, g(X_1,R_1,\dots,X_n,R_n) = \int_{(\mathbb{R}\times[0,1])^n} \mathsf{E}\, g(X_{1;p_1},R_{1;p_1},\dots,X_{n;p_n},R_{n;p_n})\, \mathrm{d}p_1 \cdots \mathrm{d}p_n,$$

where p_j and $\mathrm{d}p_j$ stand, respectively, for $x_j, \mathsf{r}_j(x_j,u_j)$ and $\mathsf{P}(X_j \in \mathrm{d}x_j)\,\mathrm{d}u_j$. Instead of the condition that g be bounded from below or above, it is enough to require only that $g(x_1, r_1, \dots, x_n, r_n) - c_1 x_1 - \dots - c_n x_n$ be so for some real constants $c_1,\dots,c_n$ over all real $x_1, r_1, \dots, x_n, r_n$.

For every natural α, let $\mathscr{H}_+^\alpha$ denote the class of all functions $f\colon \mathbb{R} \to \mathbb{R}$ such that f has finite derivatives $f^{(0)} := f, f^{(1)} := f', \dots, f^{(\alpha-1)}$ on $\mathbb{R}$, $f^{(\alpha-1)}$ is convex on $\mathbb{R}$, and $f^{(j)}(-\infty+) = 0$ for $j = 0,1,\dots,\alpha-1$.

Applying Theorem 2.4 along with results of [26; 28] to the mentioned asymmetry-corrected versions of self-normalized sums, one can obtain the following results.

Corollary 2.5. *Consider the self-normalized sum*

$$S_W := \frac{X_1 + \dots + X_n}{\frac{1}{2}\sqrt{W_1^2 + \dots + W_n^2}},$$

where $W_i := |X_i - \mathsf{r}_i(X_i,U_i)|$; here, $\frac{0}{0} := 0$. Then

$$\mathsf{E}\, f(S_W) \leqslant \mathsf{E}\, f(Z) \quad \forall f \in \mathscr{H}_+^5 \quad \text{and} \tag{2.14}$$

$$\mathsf{P}(S_W \geqslant x) \leqslant c_{5,0}\, \mathsf{P}(Z \geqslant x) \quad \forall x \in \mathbb{R}, \tag{2.15}$$

where $c_{5,0} = 5!(e/5)^5 = 5.699\dots$ and, as before, Z denotes a standard normal r.v.

Corollary 2.6. *Consider the self-normalized sum*

$$S_{Y,\lambda} := \frac{X_1 + \cdots + X_n}{(Y_1^\lambda + \cdots + Y_n^\lambda)^{\frac{1}{2\lambda}}},$$

where $Y_i := |X_i \, \mathsf{r}_i(X_i, U_i)|$. *Suppose that for some* $p \in (0,1)$ *and all* $i \in \{1, \dots, n\}$

$$\frac{X_i}{|\mathsf{r}_i(X_i, U_i)|} \, \mathrm{I}\{X_i > 0\} \leqslant \frac{1-p}{p} \text{ a.s.} \tag{2.16}$$

Then for all

$$\lambda \geqslant \lambda_*(p) := \begin{cases} \dfrac{1 + p + 2p^2}{2\big(\sqrt{p - p^2} + 2p^2\big)} & \text{if} \quad 0 < p \leqslant \frac{1}{2}, \\ 1 & \text{if} \quad \frac{1}{2} \leqslant p < 1, \end{cases} \tag{2.17}$$

one has

$$\mathsf{E}\, f(V_{Y,\lambda}) \leqslant \mathsf{E}\, f(T_n) \quad \forall f \in \mathscr{H}_+^3 \quad \text{and}$$
$$\mathsf{P}(V_{Y,\lambda} \geqslant x) \leqslant c_{3,0}\, \mathsf{P}^{\mathsf{LC}}(T_n \geqslant x) \quad \forall x \in \mathbb{R},$$

where $T_n := (Z_1 + \cdots + Z_n)/n^{1/(2\lambda)}$; $Z_1, \dots, Z_n$ *are independent r.v.'s each having the standardized Bernoulli distribution with parameter* p; *the function* $x \mapsto \mathsf{P}^{\mathsf{LC}}(T_n \geqslant x)$ *is the least log-concave majorant of the function* $x \mapsto \mathsf{P}(T_n \geqslant x)$ *on* $\mathbb{R}$; $c_{3,0} = 2e^3/9 = 4.4634\dots$. *The upper bound* $c_{3,0}\, \mathsf{P}^{\mathsf{LC}}(T_n \geqslant x)$ *can be replaced by somewhat better ones, in accordance with [25, Theorem 2.3] or [28, Corollary 4]. The lower bound* $\lambda_*(p)$ *on* λ *given by* (2.17) *is the best possible one, for each* p.

The bounded-asymmetry condition (2.16) is likely to hold when the X_i's are bounded i.i.d. r.v.'s. For instance, (2.16) holds with $p = \frac{1}{3}$ for r.v. X in Example 2.3 in place of X_i.

3 Statements of related results, with discussion

We begin this section with a number of propositions, collected in Subsections 3.1. These propositions describe general properties of the reciprocating function r and the associated functions x_+ and x_-, and thus play a dual role. On the one hand, these properties of r and $x_\pm$ may be of independent interest, each to its own extent. On the other hand, they will be used in the proofs of the basic Theorem 2.2 and related results to be stated and discussed in Subsections 3.2–3.5.

In Subsection 3.2, a generalization and various specializations of the mentioned two-point zero-mean disintegration are presented; methods of proofs are discussed and numerous relations of these results between themselves and with the mentioned result by Aizenman *et al.* [3] are also given. In Subsection 3.3, which exploits some of the results of Subsection 3.2, the disintegration based on the reciprocating function is shown to be optimal – most symmetric, but also most inhomogeneous in the widths. In Subsection 3.4, various characterizations of the reciprocating function r (as well as of the functions $x_\pm$) are given. These characterizations are perhaps the most difficult results in this paper to obtain. They are then used in Subsection 3.5 for modeling.

In all these results, the case when $X = 0$ a.s. is trivial. So, henceforth let us assume by default that $\mathsf{P}(X = 0) < 1$. Also, unless specified otherwise, μ will stand for the distribution μ_X of X.

3.1 General properties of the functions $x_\pm$ and r

Let us begin this subsection by stating, for easy reference, some elementary properties of the functions $x_\pm$ defined by (2.3) and (2.4).

Proposition 3.1. *Take any $h \in [0, m]$ and $x \in [-\infty, \infty]$. Then*

$$x \geqslant x_+(h) \iff x \geqslant 0 \ \&\ G(x) \geqslant h; \tag{3.1}$$
$$x \leqslant x_-(h) \iff x \leqslant 0 \ \&\ G(x) \geqslant h. \tag{3.2}$$

It follows that

$$G(x) < h \text{ for all } x \in [0, x_+(h)); \tag{3.3}$$
$$G(x_+(h)-) \leqslant h \leqslant G(x_+(h)); \tag{3.4}$$
$$G(x) < h \text{ for all } x \in (x_-(h), 0]; \tag{3.5}$$
$$G(x_-(h)+) \leqslant h \leqslant G(x_-(h)). \tag{3.6}$$

Moreover, for any h_1, h_2, and x one has the following implications:

$$\big(0 \leqslant h_1 < h_2 \ \&\ x_+(h_1) = x_+(h_2) = x\big) \implies \big(\mu(\{x\}) > 0 \ \&\ x > 0\big); \tag{3.7}$$
$$\big(0 \leqslant h_1 < h_2 \ \&\ x_-(h_1) = x_-(h_2) = x\big) \implies \big(\mu(\{x\}) > 0 \ \&\ x < 0\big). \tag{3.8}$$

Furthermore, the functions x_+ and $-x_-$ are

(i) non-decreasing on $[0, m]$;

(ii) finite on $[0, m)$;

(iii) strictly positive on $(0, m]$;

(iv) left-continuous on $(0, m]$.

Consider the lexicographic order $\prec$ on $[0, \infty] \times [0, 1]$ defined by the formula

$$(x_1, u_1) \prec (x_2, u_2) \iff \big(x_1 < x_2 \text{ or } (x_1 = x_2 \ \&\ u_1 < u_2)\big) \tag{3.9}$$

for all (x_1, u_1) and (x_2, u_2) in $[0, \infty] \times [0, 1]$. Extend this order symmetrically to $[-\infty, 0] \times [0, 1]$ by the formula

$$(x_1, u_1) \prec (x_2, u_2) \iff (-x_1, u_1) \prec (-x_2, u_2)$$

for all (x_1, u_1) and (x_2, u_2) in $[-\infty, 0] \times [0, 1]$.

Proposition 3.2. *The function $\tilde{G}$ is $\prec$-nondecreasing on $[0, \infty] \times [0, 1]$: if (x_1, u_1) and (x_2, u_2) are in $[0, \infty] \times [0, 1]$ and $(x_1, u_1) \prec (x_2, u_2)$, then $\tilde{G}(x_1, u_1) \leqslant \tilde{G}(x_2, u_2)$. Similarly, $\tilde{G}$ is $\prec$-nondecreasing on $[-\infty, 0] \times [0, 1]$.*

Proposition 3.3. *For all $h \in [0, m]$ (recall definition (2.7)), one has*

$$H_+(h) := \mathsf{E} X\, \mathrm{I}\{X > 0,\ \tilde{G}(X, U) \leqslant h\} = h, \tag{3.10}$$
$$H_-(h) := \mathsf{E}(-X)\, \mathrm{I}\{X < 0,\ \tilde{G}(X, U) \leqslant h\} = h. \tag{3.11}$$

The following proposition is a useful corollary of Proposition 3.3.

Proposition 3.4. *One has* $\mathsf{P}\big(X \neq 0,\ \tilde{G}(X,U) = h\big) = 0$ *for all real* h. *Therefore,* $\mathsf{P}\big(\tilde{G}(X,U) = h\big) = 0$ *for all real* $h \neq 0$*; that is, the distribution of the "randomized" version* $\tilde{G}(X,U)$ *of* $G(X)$ *may have an atom only at* 0.

Along with the r.v. X, let Y, Y_+, Y_- stand for any r.v.'s which are independent of U and whose distributions are determined by the formulas

$$\mathsf{P}(Y \in A) = \frac{\mathsf{E}\,|X|\,\mathrm{I}\{X \in A\}}{\mathsf{E}\,|X|} \quad \text{and} \quad \mathsf{P}(Y_\pm \in A) = \frac{\mathsf{E}\,|X^\pm|\,\mathrm{I}\{X \in A\}}{\mathsf{E}\,|X^\pm|} \tag{3.12}$$

for all $A \in \mathscr{B}(\mathbb{R})$; this is equivalent to

$$\mathsf{E}\,f(Y,U) = \frac{1}{2m}\,\mathsf{E}\,|X|\,f(X,U) \quad \text{and} \quad \mathsf{E}\,f(Y_\pm,U) = \frac{1}{m}\,\mathsf{E}\,|X^\pm|\,f(X,U) \tag{3.13}$$

for all Borel functions $f : \mathbb{R}^2 \to \mathbb{R}$ bounded from below (or from above). Here and elsewhere, we use the standard notation $x^+ := \max(0,x)$ and $x^- := \min(0,x)$. One should not confuse $Y_\pm$ with $Y^\pm$; in particular, by (3.12), $\mathsf{P}(Y^+ = 0) = \mathsf{P}(Y \leqslant 0) = \mathsf{P}(X \leqslant 0) \neq 0$ (since $\mathsf{E}X = 0$), while $\mathsf{P}(Y_+ = 0) = 0$.

Now one can state another corollary of Proposition 3.3:

Proposition 3.5. *One has* $\mathsf{P}(Y_+ = 0) = \mathsf{P}(Y_- = 0) = \mathsf{P}(Y = 0) = 0$ *and* $\mathsf{P}\big(\tilde{G}(Y_+,U) \leqslant h\big) = \mathsf{P}\big(\tilde{G}(Y_-,U) \leqslant h\big) = \mathsf{P}\big(\tilde{G}(Y,U) \leqslant h\big) = \frac{h}{m}$ *for all* $h \in [0,m]$. *That is, the distribution of each of the three r.v's* $\tilde{G}(Y_+,U)$, $\tilde{G}(Y_-,U)$, *and* $\tilde{G}(Y,U)$ *is uniform on the interval* $[0,m]$.

At this point one is ready to admit that the very formulation of Theorem 2.2 may seem problematic for the following reasons. On the one hand, the two-value zero-mean r.v.'s $X_{a,b}$ are not defined (and cannot be reasonably defined) when one of the points a, b is ∞ or $-\infty$ while the other one is nonzero. On the other hand, $\mathsf{r}(x,u)$ may take infinite values for some $u \in [0,1]$ and real nonzero x, which will make the r.v. $X_{x,\mathsf{r}(x,u)}$ undefined. For example, if X has the zero-mean distribution (say μ_{Exp}) with density $e^{x-1}\,\mathrm{I}\{x < 1\}$, then $\mathsf{r}(x,u) = -\infty$ for all $(x,u) \in [1,\infty) \times [0,1]$; or, if X has the distribution $\frac{1}{2}\mu_{\text{Exp}} + \frac{1}{4}\delta_{-1} + \frac{1}{4}\delta_1$, then $\mathsf{r}(x,u) = -\infty$ for $(x,u) \in \{(1,1)\} \cup \big((1,\infty) \times [0,1]\big)$.

However, such concerns are taken care of by another corollary of Proposition 3.3:

Proposition 3.6. *Almost surely,* $|\mathsf{r}(X,U)| < \infty$.

An application of Proposition 3.4 is the following refinement of Proposition 3.6. Let, as usual, $\operatorname{supp}\nu$ denote the support of a given nonnegative measure ν, which is defined as the set of all points $x \in \mathbb{R}$ such that for any open neighborhood O of x one has $\nu(O) > 0$. Then, also as usual, $\operatorname{supp}X$ is defined as the support of the distribution μ_X of X.

Proposition 3.7. *One has* $\mathsf{P}\big(X \neq 0,\ \mathsf{r}(X,U) \notin (\operatorname{supp}X) \setminus \{0\}\big) = 0$*; that is, almost surely on the event* $X \neq 0$, *the values of the r.v.* $\mathsf{r}(X,U)$ *are nonzero and belong to* $\operatorname{supp}X$. *In particular,* $\mathsf{P}\big(X \neq 0,\ \mathsf{r}(X,U) = 0\big) = 0$. *(Obviously,* $\mathsf{r}(X,U) = 0$ *on the event* $\{X = 0\}$.*)*

In the sequel, the following definition will be quite helpful:

$$\hat{x}(x,u) := \begin{cases} x_+(\tilde{G}(x,u)) & \text{if } x \in [0,\infty], \\ x_-(\tilde{G}(x,u)) & \text{if } x \in [-\infty,0] \end{cases} \tag{3.14}$$

for $u \in [0,1]$; cf. definition (2.6) of the reciprocating function r.

Proposition 3.8. *Take any or* $(x,u) \in [-\infty,\infty] \times [0,1]$ *and let* $h := \tilde{G}(x,u)$ *and, for brevity,* $\hat{x} := \hat{x}(x,u)$. *Let* μ *stand for the distribution of* X. *Then*

(i) $0 \leqslant \hat{x} \leqslant x$ *if* $x \geqslant 0$;

(ii) if $0 \leqslant \hat{x} < x$, *then all of the following conditions must occur:*

(a) $x_+(h+) > x_+(h)$;
(b) $G(\hat{x}) = G(x-) = \tilde{G}(x,u) = h$;
(c) $\mu\big((\hat{x},x)\big) = 0$;
(d) $u = 0$ *or* $\mu\big((\hat{x},x]\big) = 0$;
(e) $u = 0$ *or* $G(\hat{x}) = G(x) = h$;
(f) $u = 0$ *or* $x \neq x_+(h_1)$ *for any* $h_1 \in [0,m]$;

(iii) $0 \geqslant \hat{x} \geqslant x$ *if* $x \leqslant 0$;

(iv) if $0 \geqslant \hat{x} > x$, *then all of the following conditions must occur:*

(a) $x_-(h+) < x_-(h)$;
(b) $G(\hat{x}) = G(x+) = \tilde{G}(x,u) = h$;
(c) $\mu\big((x,\hat{x})\big) = 0$;
(d) $u = 0$ *or* $\mu\big([x,\hat{x})\big) = 0$;
(e) $u = 0$ *or* $G(\hat{x}) = G(x) = h$;
(f) $u = 0$ *or* $x \neq x_-(h_1)$ *for any* $h_1 \in [0,m]$;

(v) if $x = x_+(h_1)$ *or* $x = x_-(h_1)$ *for some* $h_1 \in [0,m]$, *then* $\hat{x}(x,u) = x$ *for all* $u \in (0,1]$.

From Proposition 3.8, we shall deduce

Proposition 3.9. *Almost surely,* $\hat{x}(X,U) = X$.

In view of Propositions 3.9 and 3.8, one may find it appropriate to refer to $\hat{x}(x,u)$ as the *regularized* version of x, and to the function $\hat{x}$ as the *regularizing function* for (the distribution of) X.

We shall use Proposition 3.9 to show that the mentioned in Introduction symmetry property $\mathsf{r}(-x) \equiv \mathsf{r}(x)$ of the reciprocating function for symmetric r.v. X with a continuous strictly increasing d.f. essentially holds in general, without the latter two restrictions on the d.f.:

Proposition 3.10. *The following conditions are equivalent to one another:*

(i) X *is symmetric;*

(ii) G *is even;*

(iii) $x_- = -x_+$;

(iv) $\mathsf{r} = -\hat{x}$;

(v) $\mathsf{r}(X,U) = -X$ *a.s.*

Propositions 3.4 and 3.9 can also be used to show that the term "reciprocating function" remains appropriate even when the d.f. of X is not necessarily strictly increasing. Toward that end, let us first state

Proposition 3.11. *For any given* $(x,u)\in\mathbb{R}\times[0,1]$*, let*

$$v:=\mathsf{v}(x,u):=\begin{cases}\frac{h-G(y+)}{G(y)-G(y+)} & \text{if } x\geqslant 0 \ \&\ G(y)\neq G(y+),\\ \frac{h-G(y-)}{G(y)-G(y-)} & \text{if } x\leqslant 0 \ \&\ G(y)\neq G(y-),\\ 1 & \text{otherwise,}\end{cases}$$

where $h:=\tilde{G}(x,u)$ *and* $y:=\mathsf{r}(x,u)$*; then*

$$\mathsf{r}\big(\mathsf{r}(x,u),v\big)=\hat{x}(x,u). \tag{3.15}$$

Moreover, the function v *is Borel and takes its values in the interval* $[0,1]$.

Now one is ready for

Proposition 3.12. *There exists a r.v.* V *taking its values in* $[0,1]$ *(and possibly dependent on* (X,U)*) such that* $\mathsf{r}\big(\mathsf{r}(X,U),V\big)=X$ *a.s. In particular, for any continuous r.v.* X *one has* $\mathsf{r}(\mathsf{r}(X))=X$ *a.s. (recall here part* (ii) *of Remark 2.1).*

Remark. In general, the identity $\mathsf{r}(x,u)=-x$ for a symmetric r.v. X does not have to hold for all $x\in\mathbb{R}$ and $u\in[0,1]$, even if X is continuous. For example, let X be uniformly distributed on $[-1,1]$ and $x>1$; then $\mathsf{r}(x,u)=\mathsf{r}(x)=-1\neq -x$ for all u. Moreover, then $\mathsf{r}(\mathsf{r}(x))=1\neq x$, so that the identity $\mathsf{r}(\mathsf{r}(x))=x$ does not have to hold for all $x\in\mathbb{R}$, even if X is continuous. Furthermore, if X is not continuous and V is not allowed to depend on (X,U), then the conclusion $\mathsf{r}\big(\mathsf{r}(X,U),V\big)=X$ a.s. in Proposition 3.12 will not hold in general. For instance, in Example 2.3 one has $\mathsf{r}\big(\mathsf{r}(1,u),v\big)=\mathsf{r}\big(\mathsf{r}(2,u),v\big)=\mathrm{I}\{v\leqslant\frac{3}{5}\}+2\,\mathrm{I}\{v>\frac{3}{5}\}$ for all u and v in $[0,1]$; so, for any r.v. V taking its values in $[0,1]$ and independent of (X,U), one has $\mathsf{P}\left(\mathsf{r}\big(\mathsf{r}(X,U),V\big)\neq X\right)\geqslant\frac{3}{10}\,\mathsf{P}(V\leqslant\frac{3}{5})+\frac{1}{10}\,\mathsf{P}(V>\frac{3}{5})\geqslant\frac{1}{10}>0$.

3.2 Variations on the disintegration theme

In this subsection we shall consider a formal extension of Theorem 2.2, stated as Proposition 3.13, which is in fact equivalent to Theorem 2.2, and yet is more convenient in certain applications. A number of propositions which are corollaries to Theorem 2.2 or Proposition 3.13 will be considered here, including certain identities for the joint distribution of X and $\mathsf{r}(X,U)$. As noted before, Theorem 2.2 implies a certain disintegration of the zero-mean distribution of X into a mixture of two-point zero-mean distributions $\big($recall (2.13)$\big)$. We shall prove that such a disintegration can be obtained directly as well, and that proof is much simpler than the proof of Theorem 2.2.

Let us now proceed by noting first a special case of (2.12), with $g(x,r):=\mathrm{I}\{x=0,r\neq 0\}$ for all real x and r. Then it follows that $\mathsf{r}(X,U)\neq 0$ almost surely on the event $\{X\neq 0\}$:

$$\mathsf{P}\left(X\neq 0,\ \mathsf{r}(X,U)=0\right)=0, \tag{3.16}$$

since $\mathsf{P}\left(X_{a,b}\neq 0,\ R_{a,b}=0\right)=0$ for any a and b with $ab\leqslant 0$. $\big($In fact, (3.16) is part of Proposition 3.7, which will be proved in Subsection 4.1 – of course, without relying on (2.12) – and then used in the proof Theorem 2.2.$\big)$ Since $\mathsf{r}(x,u)=0$ if $x=0$, (3.16) can be rewritten in the symmetric form, as

$$\mathsf{P}\big(X\,\mathsf{r}(X,U)<0 \text{ or } X=\mathsf{r}(X,U)=0\big)=1. \tag{3.17}$$

Next, note that the formalization of (2.11) given in Theorem 2.2 differs somewhat from the way in which the notion of the conditional distribution is usually understood. Yet, Theorem 2.2 and its extension, Theorem 2.4, are quite convenient in the applications, such as Corollaries 2.5 and 2.6, and others. However, Theorem 2.2 can be presented in a more general form – as a statement on the joint distribution of the ordered pair $\big(X,\mathsf{r}(X,U)\big)$ and the (unordered) set $\{X,\mathsf{r}(X,U)\}$, which may appear to be in better accordance with informal statement (2.11):

Proposition 3.13. *Let $g\colon\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ be any Borel function bounded from below (or from above), which is symmetric in the pair $(\tilde{x},\tilde{r})$ of its last two arguments:*

$$g(x,r;\tilde{r},\tilde{x})=g(x,r;\tilde{x},\tilde{r}) \tag{3.18}$$

for all real $x,r,\tilde{x},\tilde{r}$. Then

$$\begin{aligned}&\mathsf{E}\,g\big(X,\mathsf{r}(X,U);X,\mathsf{r}(X,U)\big)\\&\qquad=\int_{\mathbb{R}\times[0,1]}\mathsf{E}\,g\big(X_{x,\mathsf{r}(x,u)},R_{x,\mathsf{r}(x,u)};x,\mathsf{r}(x,u)\big)\;\mathsf{P}(X\in\mathrm{d}x)\,\mathrm{d}u.\end{aligned}$$

Instead of the condition that g be bounded from below or above, it is enough to require only that $g(x,r;\tilde{x},\tilde{r})-cx$ be so for some real constant c – over all real $x,r,\tilde{x},\tilde{r}$.

Symmetry restriction (3.18) imposed on the functions g in Proposition 3.13 corresponds to the fact that the conditioning in (2.10) and (2.11) is on the (unordered) set $\{X,\mathsf{r}(X,U)\}$, and of course not on the ordered pair $\big(X,\mathsf{r}(X,U)\big)$. Indeed, the natural conditions $\psi(a,b)=\psi(b,a)=\tilde{\psi}(\{a,b\})$ (for all real a and b) establish a one-to-one correspondence between the symmetric functions $(a,b)\mapsto\psi(a,b)$ of the ordered pairs (a,b) and the functions $\{a,b\}\mapsto\tilde{\psi}(\{a,b\})$ of the sets $\{a,b\}$. This correspondence can be used to *define* the Borel σ-algebra on the set of all sets of the form $\{a,b\}$ with real a and b as the σ-algebra generated by all symmetric Borel functions on $\mathbb{R}^2$. It is then with respect to this σ-algebra that the conditioning in the informal equation (2.11) should be understood.

Even if more cumbersome than Theorem 2.2, Proposition 3.13 will sometimes be more convenient to use. We shall prove Proposition 3.13 (later in Section 4) and then simply note that Theorem 2.2 is a special case of Proposition 3.13.

Alternatively, one could first prove Theorem 2.2 – in a virtually the same way as Proposition 3.13 is proved in this paper $\big($one only would have to use $g(a,b)$ instead of $g(a,b;a,b)[=g(a,b;b,a)]\big)$, and then it would be easy to deduce the ostensibly more general Proposition 3.13 from Theorem 2.2, in view of (3.17). Indeed, for any function g as in Proposition 3.13, one can observe that $\mathsf{E}\,g\big(X,\mathsf{r}(X,U);X,\mathsf{r}(X,U)\big)=\mathsf{E}\,\tilde{g}\big(X,\mathsf{r}(X,U)\big)$ and $\mathsf{E}\,g\big(X_{a,b},R_{a,b};a,b\big)=\mathsf{E}\,\tilde{g}\big(X_{a,b},R_{a,b}\big)$ for all real a and b such that either $ab<0$ or $a=b=0$, where $\tilde{g}(a,b):=g(a,b;a,b)$.

The following proposition, convenient in some applications, is a corollary of Proposition 3.13.

Proposition 3.14. *Let $g:=g_1-g_2$, where $g_i\colon\mathbb{R}^2\times\mathbb{R}^2\to\mathbb{R}$ ($i=1,2$) are any Borel functions bounded from below (or from above), symmetric in their last two arguments. Suppose that*

$$\begin{gathered}g(0,0;0,0)=0;\\ g(x,r;x,r)\,r=g(r,x;r,x)\,x \text{ for all real } x \text{ and } r \text{ with } xr<0.\end{gathered} \tag{3.19}$$

Then $\mathsf{E}\, g_1\big(X,\mathsf{r}(X,U);X,\mathsf{r}(X,U)\big) = \mathsf{E}\, g_2\big(X,\mathsf{r}(X,U);X,\mathsf{r}(X,U)\big)$.

Proposition 3.14 allows one to easily obtain identities for the distribution of the ordered pair $\big(X,\mathsf{r}(X,U)\big)$ or, more generally, for the conditional distribution of $\big(X,\mathsf{r}(X,U)\big)$ given the random set $\{X,\mathsf{r}(X,U)\}$.

For instance, letting $g(x,r;\tilde{x},\tilde{r}) := x\,\psi(\tilde{x},\tilde{r})$, one obtains the following proposition, which states that the conditional expectation of X given the random set $\{X,\mathsf{r}(X,U)\}$ is zero:

$$\mathsf{E}\,\big(X \,|\, \{X,\mathsf{r}(X,U)\}\big) = 0.$$

More formally, one has

Proposition 3.15. *Suppose that* $\psi\colon \mathbb{R}^2 \to \mathbb{R}$ *is a symmetric Borel function, so that* $\psi(x,r) = \psi(r,x)$ *for all real* x *and* r. *Suppose also that the function* $(x,r) \mapsto x\,\psi(x,r)$ *is bounded on* $\mathbb{R}^2$. *Then*

$$\mathsf{E} X\psi\big(X,\mathsf{r}(X,U)\big) = 0.$$

While Proposition 3.15 is a special case of Proposition 3.14 and hence of Proposition 3.13, the general case presented in Proposition 3.13 will be shown to follow rather easily from this special case; essentially, this easiness is due to the fact that a distribution on a given two-point set is uniquely determined if the mean of the distribution is known – to be zero, say, or to be any other given value.

Looking back at (3.16), one can see that the ratio $\frac{X}{\mathsf{r}(X,U)}$ can be conventionally defined almost surely on the event $\{X \neq 0\}$; let also $\frac{X}{\mathsf{r}(X,U)} := -1$ on the event $\{X = 0\}$. Letting then $g(x,r;\tilde{r},\tilde{x}) := \big(\psi(r,x) + \psi(x,r)\,\frac{x}{r}\big)\,\mathrm{I}\{xr < 0\}\,\varphi(\tilde{x},\tilde{r})$ for all real $x,r,\tilde{x},\tilde{r}$, where ψ is any nonnegative Borel function and φ is any symmetric nonnegative Borel function, one obtains from Proposition 3.14 the identity

$$\mathsf{E}\,\psi\big(X,\mathsf{r}(X,U)\big)\,\frac{X}{\mathsf{r}(X,U)}\,\varphi\big(X,\mathsf{r}(X,U)\big) = -\,\mathsf{E}\,\psi\big(\mathsf{r}(X,U),X\big)\,\varphi\big(X,\mathsf{r}(X,U)\big). \tag{3.20}$$

In particular, letting here $\psi = 1$, one sees that the conditional expectation of $\dfrac{X}{\mathsf{r}(X,U)}$ given the two-point set $\{X,\mathsf{r}(X,U)\}$ is -1:

$$\mathsf{E}\left(\frac{X}{\mathsf{r}(X,U)}\,\middle|\,\{X,\mathsf{r}(X,U)\}\right) = -1.$$

It further follows that

$$\mathsf{E}\,\frac{X}{\mathsf{r}(X,U)} = -1. \tag{3.21}$$

On the other hand, letting $\frac{\mathsf{r}(X,U)}{X} := -1$ on the event $\{X = 0\}$, one has

Proposition 3.16. *If* X *is symmetric, then* $\mathsf{E}\,\frac{\mathsf{r}(X,U)}{X} = -1$*; otherwise,*

$$\mathsf{E}\,\frac{\mathsf{r}(X,U)}{X} < -1. \tag{3.22}$$

The contrast between (3.21) and (3.22) may appear surprising, as an ostensible absence of interchangeability between X and $\mathsf{r}(X,U)$. However, this does not mean that our construction of the reciprocating function is deficient in any sense. In fact, as mentioned before, the disintegration based on r will be shown to be optimal in a variety of senses. Also, such "non-interchangeability" of X and $\mathsf{r}(X,U)$ manifests itself even in the case of a "pure" two-point zero-mean distribution:

$$\mathsf{E}\,\frac{X_{a,b}}{R_{a,b}} = -1, \quad \mathsf{E}\,\frac{R_{a,b}}{X_{a,b}} = -1 + \frac{(a+b)^2}{ab} \tag{3.23}$$

for all a and b with $ab < 0$; recall (2.9).

The "strange" inequality $\mathsf{E}\,\frac{X}{\mathsf{r}(X,U)} \neq \mathsf{E}\,\frac{\mathsf{r}(X,U)}{X}$ (unless X is symmetric) is caused only by the use of an inappropriate averaging measure – which is the distribution of r.v. X, just one r.v. of the pair $\big(X, \mathsf{r}(X,U)\big)$ – and this choice of one r.v. over the other breaks the symmetry. Here is how this concern is properly addressed:

Proposition 3.17. *For r.v.'s Y and $Y_\pm$ described in the paragraph containing* (3.12)*,*

$$\big(Y, \mathsf{r}(Y,U)\big) \stackrel{\mathrm{D}}{=} \big(\mathsf{r}(Y,U), Y\big); \tag{3.24}$$

$$\big(Y_+, \mathsf{r}(Y_+,U)\big) \stackrel{\mathrm{D}}{=} \big(\mathsf{r}(Y_-,U), Y_-\big) \stackrel{\mathrm{D}}{=} \big(x_+(H), x_-(H)\big); \tag{3.25}$$

$$\big\{Y, \mathsf{r}(Y,U)\big\} \stackrel{\mathrm{D}}{=} \big\{Y_+, \mathsf{r}(Y_+,U)\big\} \stackrel{\mathrm{D}}{=} \big\{Y_-, \mathsf{r}(Y_-,U)\big\} \stackrel{\mathrm{D}}{=} \big\{x_+(H), x_-(H)\big\}, \tag{3.26}$$

where H is any r.v. uniformly distributed on $[0,m]$. In particular, $\mathsf{r}(Y,U) \stackrel{\mathrm{D}}{=} Y$, $\mathsf{r}(Y_+,U) \stackrel{\mathrm{D}}{=} Y_- \stackrel{\mathrm{D}}{=} x_-(H)$, $\mathsf{r}(Y_-,U) \stackrel{\mathrm{D}}{=} Y_+ \stackrel{\mathrm{D}}{=} x_+(H)$, $\frac{\mathsf{r}(Y,U)}{Y} \stackrel{\mathrm{D}}{=} \frac{Y}{\mathsf{r}(Y,U)}$, *and* $\mathsf{E}\,\frac{\mathsf{r}(Y,U)}{Y} = \mathsf{E}\,\frac{Y}{\mathsf{r}(Y,U)}\big(< -1$ *except when X is symmetric, in which case one has "$= -1$" in place of "< -1"*$\big)$*; recall that Y, Y_+, and Y_- are almost surely nonzero, by Proposition 3.5.*

Just as in Proposition 3.13 versus (2.11), the equalities in distribution of the random two-point sets in (3.26) are understood as the equalities of the expected values of (say all nonnegative Borel) *symmetric* functions of the corresponding ordered pairs of r.v.'s.

Proposition 3.17 and, especially, relations (3.25) suggest an alternative way to construct the reciprocating function r. Namely, one could start with an arbitrary r.v. H uniformly distributed in $[0,m]$ and then let $Y_\pm := x_\pm(H)$. Then, by a disintegration theorem for the joint distribution of two r.v.'s (see e.g. [12, Proposition B.1]), there exist measurable functions $\mathsf{r}_\pm$ such that $\big(Y_+, \mathsf{r}_-(Y_+,U)\big) \stackrel{\mathrm{D}}{=} (Y_+, Y_-) \stackrel{\mathrm{D}}{=} \big(\mathsf{r}_+(Y_-,U), Y_-\big)$; cf. (3.25). Finally, one would let $\mathsf{r}(y,u) := \mathsf{r}_\pm(y,u)$ if $\mp y > 0$. However, this approach appears less constructive than the one represented by (2.6) and thus will not be pursued here.

Going back to (3.20) and letting there $\varphi = 1$ and $\psi(x,r) \equiv \mathrm{I}\{(x,r) \in A\}$ for an arbitrary $A \in \mathscr{B}(\mathbb{R}^2)$, one has

$$\mu_{(R,X)}(A) = \int_A \frac{-x}{r}\,\mu_{(X,R)}(\mathrm{d}x \times \mathrm{d}r),$$

where $R := \mathsf{r}(X,U)$ and μ_Z denotes the distribution of a random point Z, with the rule $\frac{-0}{0} := 1$. This means that the distribution of the random point $\big(\mathsf{r}(X,U), X\big)$ is absolutely continuous relative to that of $\big(X, \mathsf{r}(X,U)\big)$, with the function $(x,r) \mapsto \mathrm{I}\{x = r = 0\} + \frac{-x}{r}\,\mathrm{I}\{xr < 0\}$ as a Radon-Nikodym derivative.

Specializing further, with A of the form $B \times \mathbb{R}$ for some $B \in \mathscr{B}(\mathbb{R})$, one has

$$\begin{aligned}\mathsf{P}\big(\mathsf{r}(X,U)\in B\big) &= \mathsf{E}\,\mathrm{I}\{X\in B\}\,\frac{-X}{\mathsf{r}(X,U)}\\ &= \int_{B\times[0,1]}\frac{-x}{\mathsf{r}(x,u)}\,\mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u = \int_B \mathsf{P}(X\in\mathrm{d}x)\int_0^1\frac{-x}{\mathsf{r}(x,u)}\,\mathrm{d}u,\end{aligned}$$

so that the distribution of $\mathsf{r}(X,U)$ is absolutely continuous relative to that of X, with the function $x \mapsto \int_0^1 \frac{-x}{r(x,u)}\,\mathrm{d}u$ as a Radon-Nikodym derivative.

Recall now the special case (2.13) of (2.12). In particular, identity (2.13) implies that an arbitrary zero-mean distribution can be represented as the mixture of two-point zero-mean distributions. However, such a mixture representation by itself is much easier to prove (and even to state) than Theorem 2.2. For instance, one has

Proposition 3.18. *Let $g\colon \mathbb{R} \to \mathbb{R}$ be any Borel function bounded from below (or from above) such that $g(0)=0$. Then*

$$\mathsf{E}\,g(X) = \int_0^m \mathsf{E}\,g(X_h)\,\frac{\mathrm{d}h}{\mathsf{E}X_h^+}, \tag{3.27}$$

where $X_h := X_{x_+(h),x_-(h)}$.

We shall give a very short and simple proof of Proposition 3.18 (see Proof 1 on page 701), which relies only on such elementary properties of the functions x_+ and x_- as (3.1) and (iii) of Proposition 3.1. We shall also give an alternative proof of Proposition 3.18, based on [3, Theorem 2.2] as well on some properties of the functions x_+ and x_- provided by Propositions 3.8 and 3.1 of this paper. The direct proof is a bit shorter and, in our view, simpler.

This simplicity of the proof might be explained by the observation that – while Proposition 3.18 $\big($or, for that matter, identity (2.13)$\big)$ describes the one-dimensional distribution of X (as a certain mixture) – Theorem 2.2 provides a mixture representation of the two-dimensional distribution of the pair $\big(X,\mathsf{r}(X,U)\big)$, even though the distribution of this pair is completely determined by the distribution of X. Note that the random pair $\big(X,\mathsf{r}(X,U)\big)$ is expressed in terms of the reciprocating function, which in turn depends, in a nonlinear and rather complicated manner, on the distribution of X. Another indication of the simplicity of identity (3.27) is that it $\big($in contrast with (2.12) and even with (2.13)$\big)$ does not contain the randomizing random variable U. On the other hand, an obvious advantage of disintegration (2.12) is that it admits such applications to self-normalized sums as Corollaries 2.5 and 2.6.

However, there are a number of ways to rewrite (3.27) in terms similar to those of (2.13). Towards that end, for each function g as in Proposition 3.18, introduce the function Ψ_g defined by the formula

$$\Psi_g(h) := \frac{\mathsf{E}\,g(X_h)}{\mathsf{E}(X_h)^+} \quad \text{for all } h\in(0,m). \tag{3.28}$$

Then (3.27) can be rewritten as

$$\mathsf{E}\,g(X) = m\,\mathsf{E}\,\Psi_g(H), \tag{3.29}$$

where H is any r.v. uniformly distributed on the interval $[0,m]$. One such r.v. is $m\tilde{F}(X,U)$, where $\tilde{F}(x,u) := F(x-) + u\cdot\big(F(x)-F(x-)\big)$ and F is the d.f. of X. This follows in view of

Proposition 3.19. *The r.v. $\tilde{F}(X,U)$ is uniformly distributed on the interval $[0,1]$; cf. Proposition 3.5.*

Hence, for all g as in Proposition 3.18, one has an identity similar in form to (2.13):

$$\begin{aligned}\mathsf{E}\,g(X) &= m\int_{\mathbb{R}\times[0,1]} \Psi_g\big(m\tilde{F}(x,u)\big)\ \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u\\ &= m\int_{\mathbb{R}\times[0,1]} \mathsf{E}\,g\big(X_{a_+(x,u),a_-(x,u)}\big)\frac{\mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u}{\mathsf{E}(X_{a_+(x,u),a_-(x,u)})^+},\end{aligned}$$

where $a_+(x,u)$ and $a_-(x,u)$ stand for $x_+\big(m\tilde{F}(x,u)\big)$ and $x_-\big(m\tilde{F}(x,u)\big)$, respectively.

However, more interesting mixture representations are obtained if one uses Proposition 3.5 (and also Proposition 3.9) instead of Proposition 3.19:

Proposition 3.20. *Let $g\colon \mathbb{R}\to\mathbb{R}$ is any Borel function bounded from below (or from above). Then, assuming the rule $\frac{0}{\mathsf{r}(0,u)} := -1$ for all $u\in[0,1]$, one has*

$$\mathsf{E}\,g(X) = \int_{\mathbb{R}\times[0,1]} \mathsf{E}\,g\big(X_{x,\mathsf{r}(x,u)}\big)\ \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u; \tag{3.30}$$

$$\mathsf{E}\,g(X) = \int_{\mathbb{R}\times[0,1]} \mathsf{E}\,g\big(X_{x,\mathsf{r}(x,u)}\big)\,\frac{-x}{\mathsf{r}(x,u)}\ \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u; \tag{3.31}$$

$$\mathsf{E}\,g(X) = \frac{1}{2}\int_{\mathbb{R}\times[0,1]} \mathsf{E}\,g\big(X_{x,\mathsf{r}(x,u)}\big)\left(1-\frac{x}{\mathsf{r}(x,u)}\right)\ \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u. \tag{3.32}$$

Going back to (3.20) and letting therein $\psi(x,r)\equiv \mathsf{E}\,g(X_{x,r})$ and $\varphi=1$, one can rewrite the right-hand side of identity (3.31) as $\mathsf{E}\,\psi(X,R)\frac{-X}{R} = \mathsf{E}\,\psi(R,X) = \int_{\mathbb{R}^2}\psi(r,x)\,\mu_{(X,R)}(\mathrm{d}x,\mathrm{d}r) = \int_{\mathbb{R}^2}\mathsf{E}\,g(X_{r,x})\,\mu_{(X,R)}(\mathrm{d}x,\mathrm{d}r)$, so that (3.31) can be rewritten as

$$\mathsf{E}\,g(X) = \int_{\mathbb{R}^2}\mathsf{E}\,g(X_{r,x})\,\mu_{(X,R)}(\mathrm{d}x,\mathrm{d}r);$$

here, as before, $R:=\mathsf{r}(X,U)$. Similarly $\big($but in a simpler way, without using (3.20)$\big)$, identity (3.30) can be rewritten as

$$\mathsf{E}\,g(X) = \int_{\mathbb{R}^2}\mathsf{E}\,g(X_{x,r})\,\mu_{(X,R)}(\mathrm{d}x,\mathrm{d}r).$$

Now it is immediately clear why the right-hand sides of (3.30) and (3.31) are identical to each other: because $X_{x,r}\overset{\mathrm{D}}{=}X_{r,x}$. This is another way to derive (3.31): from (3.30) and (3.20). Of course, identity (3.30) is the same as (2.13), which was obtained as a special case of (2.12). Here, the point is that identity (2.13) can be alternatively deduced from the simple – to state and to prove – identity (3.27).

However, no simple way is seen to deduce (2.12) from (3.27). Toward such an end, one might start with the obvious identity $\mathsf{E}\,g\big(X,\mathsf{r}(X,U)\big) = \mathsf{E}\,g_1(X)$, where $g_1(x) := \int_0^1 g\big(x,\mathsf{r}(x,v)\big)\,\mathrm{d}v$. Then one might try to use (3.30) with g_1 in place of g, which yields

$$\mathsf{E}\,g\big(X,\mathsf{r}(X,U)\big) = \int_{\mathbb{R}\times[0,1]^2}\mathsf{E}\,g\big(X_{x,\mathsf{r}(x,u)},\mathsf{r}\big(X_{x,\mathsf{r}(x,u)},v\big)\big)\ \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u\,\mathrm{d}v.$$

At that, $\mathsf{E}\, g\big(X_{x,\mathsf{r}(x,u)}, \mathsf{r}(X_{x,\mathsf{r}(x,u)}, v)\big) = \frac{g\left(x,\mathsf{r}(x,v)\right)\mathsf{r}(x,u) - g\left(\mathsf{r}(x,u),\mathsf{r}\left(\mathsf{r}(x,u),v\right)\right)x}{\mathsf{r}(x,u)-x}$. From this, one would be able to get (2.12) if one could replace here the terms $\mathsf{r}(x,v)$ and $\mathsf{r}\big(\mathsf{r}(x,u),v\big)$ by $\mathsf{r}(x,u)$ and x, respectively, and it is not clear how this could be easily done, unless the distribution of X is non-atomic (cf. Propositions 3.11 and 3.12). Anyway, such an alternative proof would hardly be simpler than the proof of disintegration (2.12) given in this paper.

3.3 Optimality properties of the two-point disintegration

Two-value zero-mean disintegration is not unique. For example, consider the symmetric distribution $\frac{1}{10}\delta_{-2} + \frac{4}{10}\delta_{-1} + \frac{4}{10}\delta_1 + \frac{1}{10}\delta_2$ (cf. Example 2.3). This distribution can be represented either as the mixture $\frac{3}{10}(\frac{1}{3}\delta_{-2} + \frac{2}{3}\delta_1) + \frac{3}{10}(\frac{1}{3}\delta_2 + \frac{2}{3}\delta_{-1}) + \frac{4}{10}(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)$ of two asymmetric and one symmetric two-point zero-mean distributions or as the mixture $\frac{1}{5}(\frac{1}{2}\delta_{-2} + \frac{1}{2}\delta_2) + \frac{4}{5}(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)$ of two symmetric two-point zero-mean distributions; the latter representation is a special case of (2.13) or, equivalently, (3.27).

We shall show that, in a variety of senses (indexed by the continuous superadditive functions as described below), representation (2.13) of an arbitrary zero-mean distribution as the mixture of two-point zero-mean distributions is on an average most symmetric. The proof of this optimality property is based on the stated below variants of a well-known theorem on optimal transportation of mass, which are most convenient for our purposes; cf. e.g. [16] (translated in [18, pp. 57–107]), [7], [33], [29]. We need to introduce some definitions.

Let I_1 and I_2 be intervals on the real line. A function $k\colon I_1 \times I_2 \to \mathbb{R}$ is called *superadditive* if

$$k(a,c) + k(b,d) \geqslant k(a,d) + k(b,c)$$

for all a,b in I_1 and c,d in I_2 such that $a < b$ and $c < d$. So, superadditive functions are like the distribution functions on $\mathbb{R}^2$. For a function $k\colon I_1 \times I_2 \to \mathbb{R}$ to be superadditive, it is enough that it be continuous on $I_1 \times I_2$ and twice continuously differentiable in the interior of $I_1 \times I_2$ with a nonnegative second mixed partial derivative.

Let X_1 and X_2 be any r.v.'s with values in the intervals I_1 and I_2, respectively. Let

$$\tilde{X}_1 := \tilde{x}_1(H) \quad \text{and} \quad \tilde{X}_2 := \tilde{x}_2(H), \tag{3.33}$$

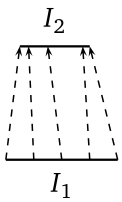
where H is any non-atomic r.v., and $\tilde{x}_1\colon \mathbb{R} \to I_1$ and $\tilde{x}_2\colon \mathbb{R} \to I_2$ are any nondecreasing left-continuous functions such that

$$\tilde{X}_1 \overset{\mathrm{D}}{=} X_1 \quad \text{and} \quad \tilde{X}_2 \overset{\mathrm{D}}{=} X_2. \tag{3.34}$$

Proposition 3.21. *Let each of the intervals I_1 and I_2 be of the form $[a,b)$, where $-\infty < a < b \leqslant \infty$. Suppose that a function k is superadditive, right-continuous, and bounded from below on $I_1 \times I_2$. Then*

$$\mathsf{E}\, k(X_1,X_2) \leqslant \mathsf{E}\, k(\tilde{X}_1,\tilde{X}_2). \tag{3.35}$$

Proposition 3.22. *Suppose that a function k is superadditive, continuous, and bounded from above on $(0,\infty)^2$. Suppose that $X_1 > 0$ and $X_2 > 0$ a.s. Then* (3.35) *holds.*



Propositions 3.21 and 3.22 essentially mean that, if the unit-transportation cost function k is superadditive, then a costliest plan of transportation of mass distribution μ_{X_1} on interval I_1 to mass distribution μ_{X_2} on I_2 is such that no two arrows in the picture here on the left may cross over; that is, smaller (respectively, larger) values in I_1 are matched with appropriate smaller (respectively, larger) values in I_2.

Note that no integrability conditions are required in Proposition 3.21 or 3.22 except for the boundedness of k from below or above; at that, either or both of the two sides of inequality (3.35) may be infinite. Proposition 3.22 is essentially borrowed from [33, Corollary 2.2.(a)]. Now we are ready to state the optimality result.

Proposition 3.23. *Suppose that one has a two-point zero-mean mixture representation of the distribution of a zero-mean r.v. X:*

$$\mathsf{E}\, g(X) = \int_S \mathsf{E}\, g(X_{y_+(s),y_-(s)})\, \nu(\mathrm{d}s) \tag{3.36}$$

for all Borel functions $g\colon \mathbb{R}\to\mathbb{R}$ bounded from below or from above, where ν is a probability measure on a measurable space (S,Σ), and $y_+\colon S\to(0,\infty)$ and $y_-\colon S\to(-\infty,0)$ are Σ-measurable functions. Then

(i) equation

$$\tilde\nu(\mathrm{d}s) := \frac{\mathsf{E} X^+_{y_+(s),y_-(s)}}{m}\,\nu(\mathrm{d}s) \tag{3.37}$$

defines a probability measure $\tilde\nu$ on (S,Σ), so that the functions y_+ and y_- can (and will be) considered as r.v.'s on the probability space $(S,\Sigma,\tilde\nu)$;

(ii) then, $y_+ \overset{\mathrm{D}}{=} Y_+$ and $y_- \overset{\mathrm{D}}{=} Y_-$, where $Y_\pm$ are r.v.'s as in (3.12);

(iii) let H be any r.v. uniformly distributed on $[0,m]$; suppose also that a superaddtive function k is either as in Proposition 3.21 $\big($with $I_1=I_2=[0,\infty)\big)$ or as in Proposition 3.22; then

$$\begin{aligned}\mathsf{E}\, k(y_+,-y_-) &\leqslant \mathsf{E}\, k\big(x_+(H),-x_-(H)\big)\\ &= \mathsf{E}\, k\big(Y_+,-\mathsf{r}(Y_+,U)\big) = \mathsf{E}\, k\big(\mathsf{r}(Y_-,U),-Y_-\big) \overset{\text{(symm)}}{=} \mathsf{E}\, k\big(\mathsf{r}(Y,U),-Y\big),\end{aligned} \tag{3.38}$$

where the symbol "$\overset{\text{(symm)}}{=}$" means an equality which takes place in the case when the additional symmetry condition $k(x,-r)=k(r,-x)$ holds for all real x and r such that $xr<0$; in particular, for any $p>0$ and $\|Z\|_p := (\mathsf{E}|Z|^p)^{1/p}$,

$$\Big\|\frac{y_\pm}{y_\mp}\Big\|_p \geqslant \Big\|\frac{x_\pm(H)}{x_\mp(H)}\Big\|_p = \Big\|\frac{Y_\pm}{\mathsf{r}(Y_\pm,U)}\Big\|_p = \Big\|\frac{\mathsf{r}(Y_\mp,U)}{Y_\mp}\Big\|_p, \tag{3.39}$$

$$\begin{aligned}\Big\|\frac{y_+}{y_-}\Big\|_p^p + \Big\|\frac{y_-}{y_+}\Big\|_p^p &\geqslant \Big\|\frac{x_+(H)}{x_-(H)}\Big\|_p^p + \Big\|\frac{x_-(H)}{x_+(H)}\Big\|_p^p\\ &= \Big\|\frac{Y_+}{\mathsf{r}(Y_+,U)}\Big\|_p^p + \Big\|\frac{\mathsf{r}(Y_+,U)}{Y_+}\Big\|_p^p = \Big\|\frac{\mathsf{r}(Y_-,U)}{Y_-}\Big\|_p^p + \Big\|\frac{Y_-}{\mathsf{r}(Y_-,U)}\Big\|_p^p\\ &= 2\Big\|\frac{\mathsf{r}(Y,U)}{Y}\Big\|_p^p = 2\Big\|\frac{Y}{\mathsf{r}(Y,U)}\Big\|_p^p;\end{aligned} \tag{3.40}$$

for any $p \geqslant 1$,

$$\begin{aligned}\|y_+ + y_-\|_p &\geqslant \big\|x_+(H)+x_-(H)\big\|_p \\ &= \big\|Y_+ + \mathsf{r}(Y_+,U)\big\|_p = \big\|\mathsf{r}(Y_-,U)+Y_-\big\|^p = \big\|\mathsf{r}(Y,U)+Y\big\|_p,\end{aligned} \tag{3.41}$$

$$\begin{aligned}\|y_+ - y_-\|_p &\leqslant \big\|x_+(H)-x_-(H)\big\|_p \\ &= \big\|Y_+ - \mathsf{r}(Y_+,U)\big\|_p = \big\|\mathsf{r}(Y_-,U)-Y_-\big\|_p = \big\|\mathsf{r}(Y,U)-Y\big\|_p;\end{aligned} \tag{3.42}$$

for any $p \leqslant 0$,

$$\begin{aligned}\mathsf{E}(y_+ - y_-)^p &\leqslant \mathsf{E}\,\big(x_+(H)-x_-(H)\big)^p \\ &= \mathsf{E}\,\big(Y_+ - \mathsf{r}(Y_+,U)\big)^p = \mathsf{E}\,\big(\mathsf{r}(Y_-,U)-Y_-\big)^p = \mathsf{E}\,\big|\mathsf{r}(Y,U)-Y\big|^p.\end{aligned} \tag{3.43}$$

Remark. Observe that the probability measure $\tilde{\nu}$ defined by (3.37) – which, according to part (ii) of Proposition 3.23, equalizes $y_\pm$ with $Y_\pm$ in distribution – is quite natural, as one considers the problem of the most symmetric disintegration of an arbitrary zero-mean distribution into the mixture of two-point zero-mean distributions as the problem of the most symmetric transportation (or, in other words, matching) of the measure $A \mapsto \mathsf{E}X^+ \mathrm{I}\{X \in A\}$ to the measure $A \mapsto \mathsf{E}(-X^-)\mathrm{I}\{X \in A\}$ (of the same total mass) or, equivalently, the most symmetric matching of the distribution of Y_+ with that of Y_-. Observe also that, in terms of $\tilde{\nu}$, mixture representation (3.36) can be rewritten in the form matching that of (3.27):

$$\mathsf{E}\,g(X) = \int_S \mathsf{E}\,g(X_{y_+(s),y_-(s)})\,\frac{m\,\tilde{\nu}(\mathrm{d}s)}{\mathsf{E}X^+_{y_+(s),y_-(s)}},$$

and at that $\int_S m\,\tilde{\nu}(\mathrm{d}s) = m = \int_0^m \mathrm{d}h$.

Remark 3.24. Inequality (3.39) means that the two-point zero-mean disintegration given in this paper is, on an average, both least-skewed to the right and least-skewed to the left. Inequality (3.39) is obtained as a special case of (3.38) (in view of Proposition 3.22) with $k(y_1,y_2) = -\frac{y_1^p}{y_2^p}$ or $k(y_1,y_2) = -\frac{y_2^p}{y_1^p}$ for positive y_1,y_2; inequality (3.40) is a "two-sided" version of (3.39). Generalizing both these one- and two-sided versions, one can take $k(y_1,y_2) \equiv -\frac{f_1(y_1)}{g_1(y_2)} - \frac{f_2(y_2)}{g_2(y_1)}$, where the functions f_1, f_2 are nonnegative, continuous, and nondecreasing, and the functions g_1, g_2 are strictly positive, continuous, and nondecreasing.

Another two-sided expression of least average skewness is given by (3.41), which is obtained as a special case of (3.38) (again in view of Proposition 3.22) with $k(y_1,y_2) \equiv -|y_1 - y_2|^p$, for positive y_1,y_2; using $-|(y_1-y_2)^\pm|^p$ instead of $-|y_1-y_2|^p$, one will have the corresponding right- and left-sided versions; note that, in any of these versions, the condition that $\|Y_+\|_p < \infty$ or $\|Y_-\|_p < \infty$ is not needed. More generally, one can take $k(y_1,y_2) \equiv -f(c_1y_1 - c_2y_2)$, where f is any nonnegative convex function and c_1, c_2 are any nonnegative constants.

On the other hand, (3.42) implies that our disintegration has the greatest p-average width $|y - \mathsf{r}(y,u)|$. This two-sided version is obtained by (3.38) in view of Proposition 3.21 with $k(y_1,y_2) \equiv |y_1+y_2|^p$, again for positive y_1,y_2; using $|(y_1+y_2)^\pm|^p$ instead will provide the corresponding right- and left-sided versions. The largest-p-average-width property can also be expressed by taking $|y_1y_2|^p$ or $|y_1^\pm y_2^\pm|^p$ in place of $|y_1+y_2|^p$ or $|(y_1+y_2)^\pm|^p$.

One can also take $k(y_1, y_2) \equiv f(c_1 y_1 + c_2 y_2)$, where f is any nonnegative convex function and c_1, c_2 are again any nonnegative constants; cf. (3.43). Thus, our disintegration can be seen as most inhomogeneous in the widths of the two-point zero-mean distributions constituting the mixture.

Another way to see this is to take any nonnegative a, b and then, in Proposition 3.23, the superadditive function $k(y_1, y_2) \equiv \mathrm{I}\{y_1 \geqslant a, y_2 \geqslant b\}$ or $k(y_1, y_2) \equiv \mathrm{I}\{y_1 < a, y_2 < b\}$. Then one sees that our disintegration makes each of the two probabilities – the large-width probability $\mathsf{P}(y_- \leqslant -a, y_+ \geqslant b)$ and the small-width probability $\mathsf{P}(-a < y_-, y_+ < b)$ – the greatest possible (over all the two-point zero-mean disintegrations, determined by the functions $y_\pm$ as in Proposition 3.23).

Moreover, *each* of these two properties – most-large-widths and most-small-widths – is equivalent to *each* of the two least-average-skewness properties: the least-right-skewness and the least-left-skewness. Indeed, for our disintegration, the right-skewness probability $\mathsf{P}(y_- > -a, y_+ \geqslant b)$ is the least possible, since it complements the large-width probability $\mathsf{P}(y_- \leqslant -a, y_+ \geqslant b)$ to $\mathsf{P}(y_+ \geqslant b)$, and it complements the small-width probability $\mathsf{P}(-a < y_-, y_+ < b)$ to $\mathsf{P}(y_- > -a)$, and at that each of the probabilities $\mathsf{P}(y_+ \geqslant b)$ and $\mathsf{P}(y_- > -a)$ is the same over all the disintegrations – recall part (ii) of Proposition 3.23. Similarly one shows that the least left-skewness is equivalent to each of the properties: most-large-widths and most-small-widths. So, there is a rigid trade-off between average skewness and width inhomogeneity.

On the other hand, reviewing the proofs of Propositions 3.21 and 3.22 (especially, see (4.13)), one realizes that the superadditive functions k of the form $k(y_1, y_2) \equiv \mathrm{I}\{y_1 \geqslant a, y_2 \geqslant b\}$ serve as elementary building blocks; more exactly, these elementary superadditive functions (together with the functions that depend only on one of the two arguments) represent the extreme rays of the convex cone that is the set of all superadditive functions. From these elementary superadditive functions, an arbitrary superadditive function can be obtained by mixing and/or limit transition. One can now conclude that the exact equivalence between the least average skewness and the most inhomogeneous width (of a two-point zero-mean disintegration) occurs at the fundamental, elementary level.

Remark. It is rather similar (and even slightly simpler) to obtain an analogue of Proposition 3.23 for the mentioned disintegration (given in [3, Theorem 2.2]) of any probability distribution into the mixture of two-point distributions with the same skewness coefficients (but possibly with different means). In fact, a same-skewness analogue of (3.42) in the limit case $p = \infty$ was obtained in [3, Theorem 2.3]; note that the corresponding L_∞ norm of the width equals ∞ unless the support of the distribution is bounded. In this paper, we shall not further pursue the matters mentioned in this paragraph.

3.4 Characteristic properties of reciprocating functions

To model reciprocating functions, one needs to characterize them. Let us begin here with some identities which follow from Proposition 3.18:

Proposition 3.25. *One has*

$$\int_0^m \frac{\mathrm{d}h}{x_+(h)} = \mathsf{P}(X > 0); \quad \int_0^m \frac{\mathrm{d}h}{-x_-(h)} = \mathsf{P}(X < 0). \tag{3.44}$$

It is interesting that identities (3.44) together with properties (i)–(iv) of Proposition 3.1 completely characterize the functions x_+ and x_-. This allows effective modeling of asymmetry patterns of a zero-mean distribution.

Proposition 3.26. *For an arbitrary* $m \in (0,\infty)$*, let* $y_+\colon [0,m] \to [0,\infty]$ *and* $-y_-\colon [0,m] \to [0,\infty]$ *be arbitrary functions with* $y_+(0) = y_-(0) = 0$ *and properties* (i)–(iv) *of Proposition 3.1 such that* $\big($*cf.* (3.44)$\big)$

$$\int_0^m \frac{\mathrm{d}h}{y_+(h)} + \int_0^m \frac{\mathrm{d}h}{-y_-(h)} \leqslant 1. \tag{3.45}$$

Then there exists a unique zero-mean distribution for which the functions x_+ *and* x_- *coincide on* $[0,m]$ *with the given functions* y_+ *and* y_-*, respectively.*

For example, take $m = 1$ and let $x_-(h) = -c$ and $x_+(h) = \frac{c}{1-h}$ for all $h \in (0,1)$, where the constant c is chosen so that the sum of the two integrals in (3.44) be $\frac{1}{2}$. Then $\mathsf{P}(X = 0) = \frac{1}{2}$, $c = 3$, $G(x) = \mathrm{I}\{x \leqslant -3\} + (1 - \frac{3}{x})\mathrm{I}\{x \geqslant 3\}$, and $\mathsf{P}(X \leqslant x) = \frac{1}{3}\mathrm{I}\{-3 \leqslant x < 0\} + \frac{5}{6}\mathrm{I}\{0 \leqslant x < 3\} + (1 - \frac{3}{2x^2})\mathrm{I}\{x \geqslant 3\}$ for all real x.

In applications such as Corollaries 2.5 and 2.6, which are stated in terms of the reciprocating function r, it is preferable to model r (rather than the functions $x_\pm$). Toward that end, let us provide various characterizations of the reciprocating function r.

For any (nonnegative) measure μ on $\mathscr{B}(\mathbb{R})$, let μ_+ be the measure defined by the formula

$$\mu_+(A) := \mu\big(A \cap [0,\infty)\big) \tag{3.46}$$

for all $A \in \mathscr{B}(\mathbb{R})$.

For any function $r\colon [-\infty,\infty] \times [0,1] \to [-\infty,\infty]$, let r_+ denote the restriction of r to the set $[0,\infty] \times [0,1]$:

$$r_+ := r|_{[0,\infty]\times[0,1]}.$$

Proposition 3.27. *Take any function* $\mathsf{s}\colon [0,\infty] \times [0,1] \to [-\infty,\infty]$ *and any measure* $\nu\colon \mathscr{B}(\mathbb{R}) \to [0,\infty)$*. Then the following two conditions are equivalent to each other:*

(I) there exists a zero-mean probability measure μ *on* $\mathscr{B}(\mathbb{R})$ *such that* $\mu_+ = \nu$ *and* $(\mathsf{r}_\mu)_+ = \mathsf{s}$*;*

(II) all of the following conditions hold:

(a) $\mathsf{s}(0,u) = 0$ *for all* $u \in [0,1]$*;*

(b) s *is* $\prec$*-nonincreasing* $\big($*recall definition* (3.9)$\big)$*;*

(c) $\mathsf{s}(x,0)$ *is left-continuous in* $x \in (0,\infty]$*;*

(d) $\mathsf{s}(x,u)$ *is left-continuous in* $u \in (0,1]$ *for each* $x \in [0,\infty]$*;*

(e) $\nu\big((-\infty,0)\big) = 0$*;*

(f) $m_\nu := G_\nu(\infty) < \infty$*;*

(g) $\big(\nu(\{x\})=0 \;\&\; x\in(0,\infty]\big) \implies \mathsf{s}(x,1)=\mathsf{s}(x,0)$;

(h) $\Big(\nu\big((x,y)\big)=0 \;\&\; 0\leqslant x<y\leqslant\infty\Big) \implies \mathsf{s}(x,1)=\mathsf{s}(y,0)$;

(i) $\big(\tilde G_\nu(x,u)<m_\nu \;\&\; (x,u)\in[0,\infty)\times[0,1]\big) \implies \mathsf{s}(x,u)>-\infty$;

(j) $\big(\tilde G_\nu(x,u)>0 \;\&\; (x,u)\in(0,\infty]\times[0,1]\big) \implies \mathsf{s}(x,u)<0$;

(k) $\displaystyle\int_{(0,\infty)\times[0,1]}\Big(1-\frac{x}{\mathsf{s}(x,u)}\Big)\,\nu(\mathrm{d}x)\,\mathrm{d}u\leqslant 1.$

Moreover, under condition (II), the measure μ as in condition (I) is unique.

In the sequel, we shall be referring to conditions (a)–(k) listed in Proposition 3.27 as 3.27(II)(a)–(k) or as 3.27(II)(a–k). Similar references will be being made to conditions listed in other propositions.

Proposition 3.27 implies, in particular, that the set of all zero-mean probability measures μ on $\mathscr{B}(\mathbb{R})$ is "parameterized" via the one-to-one mapping

$$\mu\longleftrightarrow(\nu,\mathsf{s})=\big(\mu_+,(\mathsf{r}_\mu)_+\big);$$

also, Proposition 3.27 provides a complete description of the "parameter space" (say $\mathfrak{M}$) consisting of all such pairs $(\nu,\mathsf{s})=\big(\mu_+,(\mathsf{r}_\mu)_+\big)$.

Next, we characterize the projection of the parameter space $\mathfrak{M}$ onto the "second coordinate axis"; that is, the set of all functions s such that (ν,s) is in $\mathfrak{M}$ for some measure ν. In other words, we are now going to characterize the set of the "positive parts" r_+ of the reciprocating functions of all zero-mean probability measures μ on $\mathscr{B}(\mathbb{R})$. Toward that end, with any function $\mathsf{s}\colon[0,\infty]\times[0,1]$ associate the "level" sets

$$M_{\mathsf{s}}(z):=\{(x,u)\in[0,\infty]\times[0,1]\colon \mathsf{s}(x,u)=z\},\tag{3.47}$$

for all $z\in[-\infty,\infty]$, and also

$$\begin{aligned}a_{\mathsf{s}}&:=\sup\{x\in[0,\infty]\colon \exists u\in[0,1]\ (x,u)\in M_{\mathsf{s}}(0)\};\\ b_{\mathsf{s}}&:=\inf\{x\in[0,\infty]\colon \exists u\in[0,1]\ (x,u)\in M_{\mathsf{s}}(-\infty)\};\end{aligned}\tag{3.48}$$

here $\inf\emptyset:=\infty$ and $\sup\emptyset:=0$.

Proposition 3.28. *Take any function* $\mathsf{s}\colon[0,\infty]\times[0,1]\to[-\infty,\infty]$. *Then the following two conditions are equivalent to each other:*

(I) *there exists a zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$ such that $(\mathsf{r}_\mu)_+=\mathsf{s}$;*

(II) *conditions 3.27(II)(a–d) hold, along with these three conditions:*

(i') *the set $M_{\mathsf{s}}(-\infty)$ has one of the following three forms:*

- $\emptyset$ *or*
- $[b,\infty]\times[0,1]$ *for some* $b\in(0,\infty]$ *or*
- $\{(b,1)\}\cup\big((b,\infty]\times[0,1]\big)$ *for some* $b\in(0,\infty)$*;*

in fact, this b *necessarily coincides with* b_{s}*;*

(j') *the set* $M_{\mathsf{s}}(0)$ *has one of the following two forms:*

- $[0,a]\times[0,1]$ *for some* $a\in[0,\infty]$ *or*
- $\big([0,a)\times[0,1]\big)\cup\{(a,0)\}$ *for some* $a\in(0,\infty)$*;*

in fact, this a *necessarily coincides with* a_{s}*;*

(k') $\displaystyle\int_0^1 \frac{\mathrm{d}u}{\mathsf{s}(a_{\mathsf{s}},u)}>-\infty$ *if* $\mathsf{s}(a_{\mathsf{s}},1)\neq\mathsf{s}(a_{\mathsf{s}},0)$.

Now let us characterize those $\mathsf{s}=\mathsf{r}_+$ that determine the corresponding reciprocating function r uniquely.

Proposition 3.29. *Take any function* $\mathsf{s}\colon[0,\infty]\times[0,1]\to[-\infty,\infty]$. *Then the following two conditions are equivalent to each other:*

(I) *there exists a* unique *function* r *such that* $\mathsf{r}_+=\mathsf{s}$ *and* r *coincides with the reciprocating function* r_μ *of some zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$*;*

(II) *conditions 3.27(II)(a–d) and 3.28(II)(i'–k') hold along with this almost-strict-decrease condition:*

(u) *for any* x *and* y *such that* $0\leqslant x<y\leqslant\infty$*, one of the following three conditions must occur:*

- $\mathsf{s}(x,1)>\mathsf{s}(y,0)$ *or*
- $\mathsf{s}(x,1)=\mathsf{s}(y,0)=-\infty$ *or*
- $\mathsf{s}(x,1)=\mathsf{s}(y,0)=0$.

Moreover, if either condition (I) or (II) holds, then for any zero-mean probability measure μ *on* $\mathscr{B}(\mathbb{R})$ *such that* $(\mathsf{r}_\mu)_+=\mathsf{s}$ *one has* $\operatorname{supp}(\mu_+)=\mathbb{R}\cap[a_{\mathsf{s}},b_{\mathsf{s}}]$.

Proposition 3.29 shows that the intrinsic "cause" (that is, the "cause" expressed only in terms of the function s itself) of the possible non-uniqueness of r given s is that s may fail to satisfy the almost-strict-decrease condition 3.29(II)(u), while an extrinsic "cause" of such non-uniqueness is that the support set of the "positive" part μ_+ of μ may fail to be connected. On the other hand, the next proposition shows that another extrinsic "cause" of the possible non-uniqueness is that the "negative" part μ_- of μ may fail to be non-atomic, where μ_- is the measure defined by the formula (cf. (3.46))

$$\mu_-(A):=\mu\big(A\cap(-\infty,0)\big)$$

for all $A\in\mathscr{B}(\mathbb{R})$.

Proposition 3.30. *Take any function* $\mathsf{s}\colon[0,\infty]\times[0,1]\to[-\infty,\infty]$. *Then there exists* at most one *function* r *such that* $\mathsf{r}_+=\mathsf{s}$ *and* $\mathsf{r}=\mathsf{r}_\mu$ *for some zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$ *such that* μ_- *is non-atomic. (Of course, the same conclusion holds with* μ *in place of* μ_-*.)*

Next, let us restrict our attention to the reciprocating functions of non-atomic zero-mean probability measures. Compare the following with Proposition 3.27; at that, recall Remark 2.1(ii).

Proposition 3.31. *Take any function* $\mathsf{s}\colon [0,\infty] \to [-\infty,\infty]$ *and any* non-atomic *measure* $\nu\colon \mathscr{B}(\mathbb{R}) \to \mathbb{R}$. *Then the following two conditions are equivalent to each other:*

(I) there exists a non-atomic *zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$ *such that* $\mu_+ = \nu$ *and* $(\mathsf{r}_\mu)_+ = \mathsf{s}$*;*

(II) conditions 3.27(II)(a–c,e,f,i,j) hold $\big($*with* $\mathsf{s}(x)$ *and* $G(x)$ *in place of* $\mathsf{s}(x,u)$ *and* $\tilde{G}(x,u)\big)$*, along with conditions*

(h') $0 \leqslant x < y \leqslant \infty \implies \Big(\nu\big((x,y)\big) = 0 \iff \mathsf{s}(x) = \mathsf{s}(y)\Big)$*;*

(k") $\displaystyle\int_{(0,\infty)} \Big(1 - \frac{x}{\mathsf{s}(x)}\Big)\, \nu(\mathrm{d}x) = 1.$

Moreover, under condition (II), the measure μ *as in (I) is unique.*

The following “non-atomic” version of Propositions 3.28 and 3.29 is based in part on the well-known theorem that every non-empty closed set (say in $\mathbb{R}^d$) without isolated points is the support of some non-atomic probability measure; see e.g. [23].

Proposition 3.32. *Take any function* $\mathsf{s}\colon [0,\infty] \to [-\infty,\infty]$*. Then the following two conditions are equivalent to each other:*

(I) there exists a function r *such that* $\mathsf{r}_+ = \mathsf{s}$ *and* $\mathsf{r} = \mathsf{r}_\mu$ *for some* non-atomic *zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$*;*

(II) conditions 3.27(II)(a–c) $\big($*with* $\mathsf{s}(x)$ *in place of* $\mathsf{s}(x,u)\big)$ *hold, along with condition*

(h") $\big(0 \leqslant x < y \leqslant \infty \ \&\ \mathsf{s}(x+) < \mathsf{s}(x)\big) \implies \mathsf{s}(y) < \mathsf{s}(x+)$.

Moreover, under condition (II), the function r *as in (I) is unique.*

Now we restrict our attention further, to non-atomic zero-mean probability measures with a *connected support*. Take any a_- and a_+ such that $-\infty \leqslant a_- < 0 < a_+ \leqslant \infty$ and let $I := \mathbb{R} \cap [a_-, a_+]$. The following are the “connected support” versions of Propositions 3.31 and 3.32.

Proposition 3.33. *Take any function* $\mathsf{s}\colon [0,\infty] \to [-\infty,\infty]$ *and any non-atomic measure* $\nu\colon \mathscr{B}(\mathbb{R}) \to \mathbb{R}$. *Then the following two conditions are equivalent to each other:*

(I) there exists a non-atomic *zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$ *such that* $\operatorname{supp}\mu = I$*,* $\mu_+ = \nu$*, and* $(\mathsf{r}_\mu)_+ = \mathsf{s}$*;*

(II) conditions $\mathsf{s}(0) = 0$*,* $\nu\big((-\infty,0)\big) = 0$*,* $m_\nu = G_\nu(\infty) < \infty$*, and 3.31(II)(k") hold, along with conditions*

(b') s *is strictly decreasing on* $[0, a_+]$*;*

(c') s *is continuous on* $[0,\infty]$*;*

(h") $\operatorname{supp}\nu = I_+ := \mathbb{R} \cap [0, a_+]$*;*

(i") $\mathsf{s} = a_-$ *on* $[a_+, \infty]$.

Moreover, under condition (II), the measure μ as in (I) is unique.

Proposition 3.34. *Take any function* $\mathsf{s}\colon [0,\infty] \to [-\infty,\infty]$. *Then the following two conditions are equivalent to each other:*

(I) there exists a function r *such that* $\mathsf{r}_+ = \mathsf{s}$ *and* $\mathsf{r} = \mathsf{r}_\mu$ *for some* non-atomic *zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$ *with* $\operatorname{supp}\mu = I$*;*

(II) conditions $\mathsf{s}(0) = 0$ *and 3.33(II)(b',c',i'') hold.*

Moreover, under condition (II), the function r *as in (I) is unique.*

In contrast with Proposition 3.34, the following proposition characterizes the reciprocating functions r of non-atomic zero-mean probability measures with a connected support (rather than the "positive parts" $\mathsf{s} = \mathsf{r}_+$ of such functions r).

Proposition 3.35. *Take any function* $\mathsf{r}\colon [-\infty,\infty] \to [-\infty,\infty]$. *Then the following two conditions are equivalent to each other:*

(I) there exists a non-atomic *zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$ *such that* $\operatorname{supp}\mu = I$ *and* $\mathsf{r}_\mu = \mathsf{r}$*;*

(II) condition $\mathsf{r}(0) = 0$ *holds, along with the following:*

(b'') r *is strictly decreasing on* $[a_-, a_+]$*;*

(c'') r *is continuous on* $[-\infty,\infty]$*;*

(i''') $\mathsf{r} = a_+$ *on* $[-\infty, a_-]$ *and* $\mathsf{r} = a_-$ *on* $[a_+,\infty]$*;*

($\mathsf{r}\circ\mathsf{r}$) $\mathsf{r}\big(\mathsf{r}(x)\big) = x$ *for all* $x \in [a_-, a_+]$.

Our final characterization concerns the case when it is desirable to avoid zero-mean probability measures μ with a density that is discontinuous at 0 (say, as an unlikely shape).

Proposition 3.36. *Take any function* $\mathsf{r}\colon [-\infty,\infty] \to [-\infty,\infty]$. *Then the following two conditions are equivalent to each other:*

(I) there exists a non-atomic *zero-mean probability measure* μ *on* $\mathscr{B}(\mathbb{R})$ *such that* $\mathsf{r}_\mu = \mathsf{r}$, $\operatorname{supp}\mu = I$, *and in a neighborhood of* 0 *measure* μ *has a continuous strictly positive density;*

(II) condition 3.35(II) holds, along with the following: r *is continuously differentiable in a neighborhood of* 0.

Moreover, if either condition (I) or (II) holds, then necessarily $\mathsf{r}'(0) = -1$, *that is, one has the approximate local symmetry condition* $\mathsf{r}(x) \sim -x$ *as* $x \to 0$.

3.5 Modeling reciprocating functions

We propose modeling asymmetry using reciprocating functions. In view of Propositions 3.34 and 3.35, the reciprocating function r of any non-atomic zero-mean probability measure μ with a connected support can be constructed as follows.

Construction 1. *(i) Take any a_- and a_+ such that $-\infty \leqslant a_- < 0 < a_+ \leqslant \infty$ and let $I_+ := \mathbb{R} \cap [0, a_+]$.*

(ii) Take any function $\mathsf{s}\colon [0,\infty] \to [-\infty,\infty]$ such that $\mathsf{s}(0) = 0$, s equals a_- on $[a_+,\infty]$ and is strictly decreasing and continuous on $[0, a_+]$.

(iii) Define r by the formula

$$\mathsf{r} := \begin{cases} a_+ & \text{on } [-\infty, a_-]; \\ \tilde{\mathsf{s}}^{-1} & \text{on } [a_-, 0]; \\ \mathsf{s} & \text{on } [0,\infty], \end{cases}$$

where $\tilde{\mathsf{s}} := \mathsf{s}|_{[0,a_+]}$, the restriction of the function s to the interval $[0, a_+]$.

Example 3.37. In accordance with Construction 1 and Proposition 3.36, one can suggest the two-parameter family of reciprocating functions defined by the formula:

$$\mathsf{r}(x) := \mathsf{r}_{p,c}(x) := \begin{cases} \frac{c}{p}\left(1 - (1 + x/c)^p\right) & \text{for } x \in [0,\infty]; \\ c\left((1 - px/c)^{1/p} - 1\right) & \text{for } x \in (-\infty, 0] \text{ such that } px < c, \end{cases}$$

with the convention that $\mathsf{r}(\infty) := \mathsf{r}(\infty-)$; here, $p \in \mathbb{R} \setminus \{0\}$ and $c > 0$ are real numbers, which may be referred to as the shape and scale parameters, respectively. Indeed, one can see that mere rescaling of μ (or a corresponding r.v. X) results only in a change of c: if $\mathsf{r}_{p,1} = \mathsf{r}_X$ for some zero-mean r.v. X, then $\mathsf{r}_{p,c} = \mathsf{r}_{cX}$. For $x \in [-\infty, 0]$ such that $px \geqslant c$ (that is, for $x \in [-\infty, \frac{c}{p}]$ when $p < 0$), we set $\mathsf{r}_{p,c}(x) := \infty$, in accordance with the general description of Construction 1. Let us also extend the family of functions $\mathsf{r}_{p,c}$ to $p = 0$ by continuity:

$$\mathsf{r}_{0,c}(x) := \lim_{p\to 0} \mathsf{r}_{p,c}(x) = \begin{cases} -c \ln(1 + x/c) & \text{for } x \in [0,\infty]; \\ c\,(e^{-x/c} - 1) & \text{for } x \in [-\infty, 0]. \end{cases}$$

The corresponding intervals $[a_-, a_+]$ here coincide with $[-\infty,\infty]$ if $p \geqslant 0$ and with $[\frac{c}{p},\infty]$ if $p < 0$.

Case $p = 1$ corresponds to the pattern of perfect symmetry of μ; that is, $\mathsf{r}(x) = -x$ for all x (recall Proposition 3.10). Case $p > 1$ corresponds to a comparatively long (or, equivalently, heavy) left tail of μ, so that μ will be skewed to the left. Similarly, case $p < 1$ corresponds to a comparatively long (or heavy) right tail of μ. Thus, p can be considered as the asymmetry parameter.

Another limit case is when $p \to \pm\infty$ and $c \to \infty$ in such a manner that $\frac{c}{p} \to \pm\lambda$, for some $\lambda \in (0,\infty)$, and this limit is given by

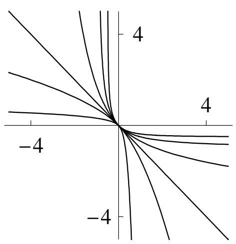
$$\mathsf{r}_{\pm\infty,\lambda}(x) := \begin{cases} \pm\,\lambda\,(1 - e^{\pm x/\lambda}) & \text{for } x \in [0,\infty]; \\ \pm\lambda \ln(1 \mp x/\lambda) & \text{for } x \in [-\infty, 0] \text{ such that } \pm x < \lambda, \end{cases}$$

where λ and $\pm\infty$ play, respectively, the roles of the scale and shape (or, more specifically, asymmetry) parameters.

Yet another limit case is when $c \to 0$ and $p \to 1$ in such a manner that $c^{p-1} \to \kappa$, for some $\kappa \in (0,\infty)$, and this limit is given by

$$\mathsf{r}_{1,0;\kappa}(x) := \begin{cases} -x/\kappa & \text{for } x \in [0,\infty]; \\ -\kappa x & \text{for } x \in [-\infty,0], \end{cases}$$

However, in this case the property $\mathsf{r}'(0) = -1$ is lost (in fact, $\mathsf{r}_{1,0;\kappa}$ is not differentiable at 0) unless $\kappa = 1$, so that, by Proposition 3.36, no corresponding zero-mean distribution μ can have a density that is strictly positive and continuous at 0.



Here on the left one can see parts of the graphs $\{(x,\mathsf{r}_{p,1}(x)): x \in [a_-,a_+]\}$ with $p = -2,-1,0,1,2,8$. Each graph is symmetric about the diagonal $\Delta := \{(x,x): x \in [-\infty,\infty]\}$, as it should be according to the reciprocity property 3.35(II)($\mathsf{r}\circ\mathsf{r}$). The tighter the graph of the reciprocating function embraces the first quadrant, the more skewed is the corresponding distribution to the right; and the tighter the graph embraces the third quadrant, the more skewed is the distribution to the left. In this example, the greater is p, the more skewed to the left must the corresponding zero-mean distribution μ be.

Construction 2. *Reciprocity property 3.35(II)($\mathsf{r}\circ\mathsf{r}$) of r implies that the graph $\{(x,\mathsf{r}(x)): x \in [a_-,a_+]\}$ can be obtained in the form $\{(x,y): F(x,y) = 0, x \in [a_-,a_+], y \in [a_-,a_+]\}$, where F is a symmetric function, which must also satisfy condition $F(0,0) = 0$, since $\mathsf{r}(0) = 0$.*

A simplest such function is the quadratic function F given by the formula

$$F(x,y) \equiv Ax^2 + 2Bxy + Ay^2 + cx + cy, \tag{3.49}$$

so that the graphs are elliptic or hyperbolic arcs symmetric about the diagonal Δ and passing through the origin. However, here we shall not consider this construction in detail.

Instead, let us turn to

Construction 3. *The symmetry of the graph $\{(x,\mathsf{r}(x)): x \in (a_-,a_+)\}$ of a reciprocating function r about the diagonal Δ suggests that r is uniquely determined by a function (say a) that maps, for each $x \in (a_-,a_+)$, the width $\mathsf{w}(x) := |x - \mathsf{r}(x)|$ to the asymmetry $\alpha(x) := x + \mathsf{r}(x)$ of the zero-mean distribution on the two-point set $\{x,\mathsf{r}(x)\}$. (Note that, by 3.35(II)(b",c"), the width function w is continuous on $[-\infty,\infty]$, strictly increasing on $[0,a_+]$ (from 0 to $a_+ - a_-$), and strictly decreasing on $[a_-,0]$ (from $a_+ - a_-$ to 0).) The function a may be referred to as the* asymmetry pattern function *of a given zero-mean distribution.*

Details of Construction 3 are presented in

Proposition 3.38.

(I) If r is a reciprocating function as in Proposition 3.35, then there exists a unique function $\mathsf{a}\colon [0,a_+ - a_-) \to \mathbb{R}$ such that $\mathsf{a}(0) = 0$,

$$x + \mathsf{r}(x) = \mathsf{a}(|x - \mathsf{r}(x)|) \tag{3.50}$$

for all $x \in (a_-, a_+)$, and the following strict Lip(1) condition (Lipschitz with constant factor 1) holds:

$$|\mathsf{a}(w_2) - \mathsf{a}(w_1)| < w_2 - w_1$$

for all w_1 and w_2 such that $0 \leqslant w_1 < w_2 < a_+ - a_-$. Also, $\mathsf{a}(w) \to a_+ + a_-$ as $w \uparrow a_+ - a_-$.

(II) *Vice versa, if a function $\mathsf{a}\colon [0, a_+ - a_-) \to \mathbb{R}$ is strictly Lip(1), $\mathsf{a}(0) = 0$, and $\mathsf{a}(w) \to a_+ + a_-$ as $w \uparrow a_+ - a_-$, then there exists a unique reciprocating function r such as in Proposition 3.35 that satisfies condition (3.50). In fact, then one necessarily has*

$$\mathsf{r}(x) = \begin{cases} -\rho\big(\xi^{-1}(x)\big) & \text{if } x \in [0, a_+); \\ \xi\big(\rho^{-1}(-x)\big) & \text{if } x \in (a_-, 0], \end{cases} \tag{3.51}$$

where the functions ξ and ρ are defined by

$$\xi(w) := \tfrac{1}{2}(w + \mathsf{a}(w)) \text{ and } \rho(w) := \tfrac{1}{2}(w - \mathsf{a}(w)) \text{ for all } w \in [0, a_+ - a_-), \tag{3.52}$$

and they are continuously and strictly increasing on $[0, a_+ - a_-)$ from 0 to a_+ and $-a_-$, respectively.

(III) *Moreover, a reciprocating function r such as in Proposition 3.35 is continuously differentiable in a neighborhood of 0 if and only if the corresponding asymmetry pattern function a is continuously differentiable in an open right neighborhood (r.n.) of 0 and $\mathsf{a}'(0+) = 0$.*

In particular, Proposition 3.38 shows that the asymmetry pattern function a is necessarily Lipschitz and hence absolutely continuous, with a density $\mathsf{a}'(w) = \frac{\mathrm{d}\mathsf{a}(w)}{\mathrm{d}w}$ such that

$$-1 < \mathsf{a}'(w) < 1$$

for almost all $w \in [0, a_+ - a_-)$. In view of (3.50), this density a' may be considered as the rate of change of asymmetry $\alpha = x + \mathsf{r}(x)$ relative to the varying width $w = |x - \mathsf{r}(x)|$ of the constituent zero-mean distribution on the two point set $\{x, \mathsf{r}(x)\}$. For instance, if at the given width w this rate $\mathsf{a}'(w)$ is close to 1, then at this width w the distribution's skewness to the right is growing fast. Also, for all $w \in (0, a_+ - a_-)$, the ratio

$$\frac{\mathsf{a}(w)}{w} = \frac{1}{w} \int_0^w \mathsf{a}'(v)\,\mathrm{d}v$$

represents the average asymmetry-to-width rate over all widths from 0 to w. Thus, Construction 3 provides a flexible and sensitive tool to model asymmetry patterns.

One can see that in Example 3.37 the asymmetry-to-width rate a' strictly increases or decreases from 0 to 1 or -1 as w increases from 0 to ∞, depending on whether $p < 1$ or $p > 1$, and $\mathsf{a}'(w) = 0$ for all $w \in [0, \infty)$ if $p = 1$. Moreover,

$$\mathsf{a}'(w) \sim \begin{cases} 1 - c_1\, e^{-w/\lambda} & \text{if } p = -\infty, \\ 1 - c_2\, w^{p-1} & \text{if } -\infty < p \leqslant 0, \\ 1 - c_3\, w^{1-1/p} & \text{if } 0 < p < 1, \\ -1 + c_4\, w^{1/p-1} & \text{if } 1 < p \leqslant \infty \end{cases}$$

as $w \to \infty$, where $c_1, \dots, c_4$ are positive real constants, depending only on the parameters p and c; here $1/p - 1 := -1$ for $p = \infty$.

Let us now provide examples of two parametric families of reciprocating functions obtained using the asymmetry-to-width rate a' as the starting point.

Example 3.39. Take any $\alpha \in [-1,1]$ and $c \in (0,\infty)$, and consider the asymmetry-to-width rate of the form

$$\mathsf{a}'(w) = \alpha\Big(1 - \frac{c^2}{(c+w)^2}\Big)$$

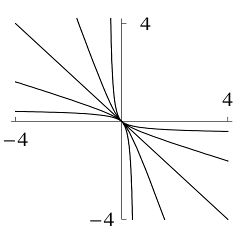
for all $w \in [0,\infty)$, so that, for $\alpha \in (0,1]$, the rate $\mathsf{a}'(w)$ increases from 0 to α as w increases from 0 to ∞; similarly, for $\alpha \in [-1,0)$, the rate $\mathsf{a}'(w)$ decreases from 0 to α as w increases from 0 to ∞. Then the corresponding asymmetry pattern function a is given by

$$\mathsf{a}(w) = \mathsf{a}_{\alpha,c}(w) = \alpha\frac{w^2}{c+w}$$

for all $w \in [0,\infty)$, and, by (3.51), the corresponding reciprocating function r is given by

$$\mathsf{r}(x) = \mathsf{r}_{\alpha,c}(x) = \frac{c + 2\alpha x - \sqrt{(c+2|x|)^2 + 8\alpha c x}}{2(\alpha + \operatorname{sign} x)}$$

for $\alpha \in (-1,1)$ and all $x \in \mathbb{R}$; expressions for $\mathsf{r}_{1,c}$ and $\mathsf{r}_{-1,c}$ are of different forms. Note that $\mathsf{r}_{\alpha,c}(x) \sim \frac{\alpha\mp1}{\alpha\pm1}x$ as $x \to \pm\infty$, for each $\alpha \in (-1,1)$; on the other hand, $\mathsf{r}_{1,c}((-\frac{c}{2})+) = \infty$, $\mathsf{r}_{1,c}(\infty-) = -\frac{c}{2}$, $\mathsf{r}_{-1,c}(\frac{c}{2}-) = -\infty$, $\mathsf{r}_{-1,c}((-\infty)+) = \frac{c}{2}$. The parameters α and c are, respectively, the shape (or, more specifically, asymmetry) and scale parameters. The graph of $\mathsf{r}_{\alpha,c}$ is the union of two hyperbolic arcs of two different hyperbolas: $-(1+\alpha)r^2 + 2\alpha x r + (1-\alpha)x^2 + cr + cx = 0$ (used for $x \geqslant 0$) and $(1-\alpha)r^2 + 2\alpha x r - (1+\alpha)x^2 + cr + cx = 0$ (used for $x \leqslant 0$) – cf. (3.49).



Yet, by Proposition 3.38, all these reciprocating functions $\mathsf{r}_{\alpha,c}$ are continuously differentiable in a neighborhood of 0 (in fact, they are so wherever on $\mathbb{R}$ they take finite values). On the left one can see parts of the graphs $\{(x, \mathsf{r}_{\alpha,1}(x)) : x \in [a_-, a_+]\}$ with $\alpha = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1$. In such an example, the shape (or, more specifically, asymmetry) parameter α can also be considered as a scale parameter – but in the direction of the diagonal $\Delta = \{(x,x) : x \in [-\infty,\infty]\}$.

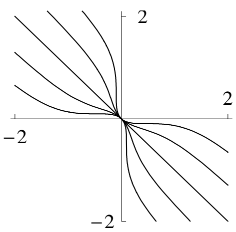
Example 3.40. Take any $\alpha \in [-1,1]$ and $c \in (0,\infty)$, and consider the asymmetry-to-width rate of the form

$$\mathsf{a}'(w) = \frac{16\alpha c^3}{3\sqrt{3}}\,\frac{w}{(c^2+w^2)^2}$$

for all $w \in [0,\infty)$, so that, for $\alpha \in (0,1]$, the rate $\mathsf{a}'(w)$ increases from 0 to α and then decreases from α to 0 as w increases from 0 to $c/\sqrt{3}$ to ∞; similarly, for $\alpha \in [-1,0)$, the rate $\mathsf{a}'(w)$ decreases from 0 to α and then increases from α to 0 as w increases from 0 to $c/\sqrt{3}$ to ∞. The corresponding asymmetry pattern function a is given by

$$\mathsf{a}(w) = \frac{8\alpha c}{3\sqrt{3}}\,\frac{w^2}{c^2+w^2}$$

for all $w \in [0,\infty)$, and, using (3.51), one can see that the corresponding reciprocating function $\mathsf{r} = \mathsf{r}_{\alpha,c}$ is given by an algebraic expression involving certain cubics. In particular, $\mathsf{r}_{\alpha,c}(x) \sim -x + \frac{8\alpha c}{3\sqrt{3}}$ as $|x| \to \infty$.



Again, the parameters α and c are, respectively, the shape (or, more specifically, asymmetry) and scale parameters. Alternatively, in this example as well, the shape/asymmetry parameter α can also be considered as a scale parameter, in the direction of the diagonal Δ. Here on the left one can see parts of the graphs $\{(x, \mathsf{r}_{\alpha,1}(x)) : x \in [a_-, a_+]\}$ with $\alpha = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1$.

Construction 4. *Looking back at Proposition 3.38, one can see that yet another way to construct an arbitrary reciprocating function* r *as in Proposition 3.35 is by using* (3.51) *with arbitrary functions* ξ *and* ρ *that are continuously and strictly increasing on* $[0, a_+ - a_-)$ *from* 0 *to* a_+ *and* $-a_-$*, respectively* (*and also using condition 3.35(II)(i''') to complete the construction of* r). *In fact, the functions* ξ *and* ρ *defined by* (3.52) *also satisfy the strict Lip(1) condition; still, even if* ξ *or* ρ *violates this Lip(1) restriction, the function* r *defined by* (3.51) *will have all the characteristic properties 3.35(II)(b'')–(*$\mathsf{r}\circ\mathsf{r}$*). However, in this paper we shall not pursue this construction further.*

Examples 3.37, 3.39, 3.40 of parametric families of reciprocating functions already appear to represent a wide enough variety. Moreover, Constructions 1–4 given in this subsection appear convenient and flexible enough for efficient modeling of asymmetry patterns that may arise in statistical practice. In any case, each of these constructions – of reciprocating functions for non-atomic distributions with connected support – is quite universal. (For discrete distributions, it appears more convenient to model asymmetry patterns based on the characterization of the functions $x_\pm$ provided by Proposition 3.26.) In any such parametric or nonparametric model, the reciprocating function can be estimated in a standard manner, as follows: substituting the empirical distribution for the "true" unknown distribution μ, one obtains empirical estimates of the function G and hence empirical estimates of the functions $x_\pm$ and r; then, if desired, the empirical estimate of r can be fit into an appropriate parametric family of reciprocating functions.

4 Proofs

In Subsection 4.1) we shall prove the propositions stated in Section 3 and then, in Subsection 4.2, the theorems and corollaries stated in Section 2.

4.1 Proofs of propositions

Proof of Proposition 3.1. Implication $\Longleftarrow$ in (3.1) follows immediately from definition (2.3), since $x_+(h)$ is a lower bound of the corresponding set. Implication $\Longrightarrow$ in (3.1) follows from (2.3) as well, taking also into account that, by (2.2), the function G is non-decreasing on $[0,\infty]$ and right-continuous on $[0,\infty)$. Thus, one obtains (3.1). Equivalence (3.2) is proved similarly.

Inequalities (3.3) and (3.5) follow immediately from (3.1) and (3.2). The first inequalities in (3.4) and (3.6) follow immediately from (3.3) and (3.5), while the second ones are special cases of (3.1) and (3.2), respectively.

Next, let us prove (3.7) and (3.8). Assume indeed that $0 \leqslant h_1 < h_2$ and $x_+(h_1) = x_+(h_2) = x$. Then, by (3.1) and (3.4), $G(x) \geqslant h_2 > h_1 \geqslant G(x-)$, so that $x\,\mu(\{x\}) > 0$. This proves (3.7). Quite similarly one proves (3.8).

Property (i) follows immediately from definitions (2.3) and (2.4).

Property (ii) follows because $G(x) \to m$ as $|x| \to \infty$.

Since the functions $\pm x_\pm$ are nonnegative $\big($by definitions (2.3) and (2.4)$\big)$ and $G(0) = 0$, property (iii) follows by (3.1) and (3.2), which imply that $G(x_\pm(h)) \geqslant h > 0$ for all $h \in (0, m]$.

Finally, let us now prove property (iv). Take any $h_0 \in (0, m]$ and let $x_0 := x_+(h_0)$. Then, by property (iii), one has $x_0 > 0$. Next, take any $x \in (0, x_0)$. Then, by (3.3) and (3.4), $G(x) < h_0 \leqslant G(x_0)$ and, by (3.1), one has $x_+(h) \in (x, x_0]$ for all h in the interval $(G(x), G(x_0)]$ and hence for all h in the nonempty subinterval $(G(x), h_0]$ of $(G(x), G(x_0)]$. This implies that x_+ is left-continuous on $(0, m]$; similarly, $-x_-$ is so. $\square$

Proof of Proposition 3.2. If $0 \leqslant x_1 < x_2 \leqslant \infty$ and $0 \leqslant u_1, u_2 \leqslant 1$ then $\tilde G(x_1, u_1) \leqslant G(x_1) \leqslant G(x_2-) \leqslant \tilde G(x_2, u_2)$; and if $0 \leqslant x \leqslant \infty$ and $0 \leqslant u_1 < u_2 \leqslant 1$ then $\tilde G(x, u_1) \leqslant \tilde G(x, u_2)$, by (2.5). This shows that the function $\tilde G$ is indeed $\prec$-nondecreasing on $[0, \infty] \times [0, 1]$. Similarly it is shown that $\tilde G$ is $\prec$-nondecreasing on $[-\infty, 0] \times [0, 1]$. $\square$

Proof of Proposition 3.3. Identity (3.11) follows from (3.10) by substituting $-X$ for X. So, it remains to prove that $H_+(h) = h$ for all $h \in [0, m]$. Fix any $h \in [0, m]$ and write

$$H_+(h) = \mathsf{E}\, X\, \mathrm{I}\{(X, U) \in M\}, \quad \text{where} \quad M := \{(x, u) \in (0, \infty] \times [0, 1] \colon \tilde G(x, u) \leqslant h\}.$$

Introduce also $x_h := \sup\{z \in [0, \infty] \colon G(z-) \leqslant h\}$. Then $x_h = \max\{z \in [0, \infty] \colon G(z-) \leqslant h\}$ $\big($because the function $z \mapsto G(z-)$ is left-continuous on $(0, \infty]\big)$. So, $G(x_h-) \leqslant h < G(z-)$ for all $z > x_h$, whence $G(x_h-) \leqslant h \leqslant G(x_h)$. Now one has to distinguish the following two cases.

Case 1: $G(x_h-) = G(x_h)$. Then $G(x_h-) = h = G(x_h)$ and $M = (0, x_h] \times [0, 1]$ $\big($because (i) $\tilde G(z, u) \leqslant G(z) \leqslant G(x_h-) = h$ for all $(z, u) \in (0, x_h) \times [0, 1]$, (ii) $\tilde G(x_h, u) = h$ for all $u \in [0, 1]$, and (iii) $\tilde G(z, u) \geqslant G(z-) > h$ for all $(z, u) \in (x_h, \infty) \times [0, 1]\big)$. It follows that $H_+(h) = \mathsf{E}\, X\, \mathrm{I}\{0 < X \leqslant x_h\} = G(x_h) = h$, whence $H_+(h) = h$ in Case 1.

Case 2: $G(x_h-) < G(x_h)$. Then $u_h := (h - G(x_h-))/(G(x_h) - G(x_h-)) \in [0, 1]$ and $\tilde G(x_h, u_h) = h$. Also, reasoning as in Case 1, here one can see that $M = \big((0, x_h) \times [0, 1]\big) \cup \big(\{x_h\} \times [0, u_h]\big)$. It follows that

$$\begin{aligned} H_+(h) &= \mathsf{E}\, X\, \mathrm{I}\{0 < X < x_h\} + x_h\, \mathsf{P}(X = x_h)\, \mathsf{P}(U \in [0, u_h]) \\ &= G(x_h-) + (G(x_h) - G(x_h-))\, u_h = \tilde G(x_h, u_h) = h, \end{aligned}$$

whence $H_+(h) = h$ in Case 2 as well. $\square$

Proof of Proposition 3.4. By Proposition 3.3, for all $h \in [0, m]$, one has $0 = H_+(h) - H_+(h-) = \mathsf{E}\, X\, \mathrm{I}\{X > 0,\ \tilde G(X, U) = h\}$, whence $\mathsf{P}(X > 0,\ \tilde G(X, U) = h) = 0$; similarly, $\mathsf{P}(X < 0, \tilde G(X, U) = h) = 0$; note also that $\tilde G(X, U) \in [0, m]$ a.s. $\square$

Proof of Proposition 3.5. This follows immediately from Proposition 3.3, since, by (3.12), $\mathsf{P}\left(\tilde G(Y_\pm, U) \leqslant h\right) = \frac{1}{m} H_\pm(h)$ and $\mathsf{P}\left(\tilde G(Y, U) \leqslant h\right) = \frac{1}{2m}\big(H_+(h) + H_-(h)\big)$ for all $h \in [0, m]$. $\square$

Proof of Proposition 3.6. This follows from Proposition 3.4. Indeed, in view of property (ii) in Proposition 3.1 and definition (2.6), the event $\{|\mathsf{r}(X,U)| = \infty\}$ is contained in the event $\{X \neq 0,\ \tilde{G}(X,U) = m\}$. ☐

Proof of Proposition 3.7. For any $h \in [0,m]$, it follows from (3.1) that $G(x) < h \leqslant G(x_+(h))$ and hence $\mathsf{P}(X \in (x, x_+(h)]) > 0$ for all $x \in [0, x_+(h))$, so that $x_+(h) \in \operatorname{supp} X$ provided that $x_+(h) \in (0,\infty)$. Similarly, $x_-(h) \in \operatorname{supp} X$ whenever $x_-(h) \in (-\infty, 0)$. So, for any $u \in (0,1)$ and $x \in \mathbb{R}$ one has $\mathsf{r}(x,u) \in \operatorname{supp} X$ whenever $\mathsf{r}(x,u) \in \mathbb{R} \setminus \{0\}$.

Next, if $\mathsf{r}(x,u) = 0$ for some $x \in \mathbb{R} \setminus \{0\}$ and $u \in (0,1)$, then $\tilde{G}(x,u) = 0$ (by (2.6) and property (iii) of Proposition 3.1). So, by Proposition 3.4, $\mathsf{P}(X \neq 0,\ \mathsf{r}(X,U) = 0) = 0$.

It remains to recall Proposition 3.6. ☐

Proof of Proposition 3.8.

(i) For $x \in [0,\infty]$, one has $G(x) \geqslant \tilde{G}(x,u) = h$, so that $\hat{x}(x,u) = x_+(h) \leqslant x$, by (2.3); also, (3.14) clearly implies that here $\hat{x}(x,u) \geqslant 0$. This proves part (i) of the proposition; part (iii) is quite similar.

(ii) Assume that $0 \leqslant \hat{x}(x,u) < x$.

(a): Note that $h = \tilde{G}(x,u) \leqslant m < \infty$. Take any $h_1 \in (h,\infty)$ and then any $z \in [0,\infty]$ such that $G(z) \geqslant h_1$ (if such a point z exists). Then $G(z) > h = \tilde{G}(x,u) \geqslant G(x-)$; since G is nondecreasing on $[0,\infty]$, it follows that $z \geqslant x$. That is, $z \geqslant x$ for all $z \in [0,\infty]$ such that $G(z) \geqslant h_1$. So, by (2.3), $x_+(h_1) \leqslant x$ for all $h_1 \in (h,\infty)$ and hence $x_+(h+) \geqslant x > \hat{x}(x,u) = x_+(h)$. This verifies condition (a) of part (ii).

(b): Using the monotonicity of G, condition $0 \leqslant \hat{x} < x$, (3.14), (3.1), and (2.5), one has $G(x-) \geqslant G(\hat{x}) \geqslant \tilde{G}(x,u) \geqslant G(x-)$. Now condition (b) of part (ii) follows.

(c): By just checked condition (b), $G(\hat{x}) = G(x-)$. Now condition (c) follows by (2.1).

(d), (e): Again by condition (b), $\tilde{G}(x,u) = G(x-)$. So, if $u \neq 0$ then, by (2.5), $G(x) = G(x-)$, whence condition (e) follows by (b). In turn, condition (d) follows from (e).

(f): Assume that $u \neq 0$ and $x = x_+(h_1)$ for some $h_1 \in [0,m]$. On the other hand, by (3.14), $\hat{x} = x_+(h)$. So, the condition $\hat{x} < x$ means that $x_+(h) < x_+(h_1)$, whence $h < h_1$, by property (i) of Proposition 3.1. Also, $x = x_+(h_1)$ implies $G(x) \geqslant h_1$, by (3.1). So, $G(x) \geqslant h_1 > h$. This contradicts condition (e) and thereby verifies condition (f).

Thus, part (ii) of the proposition is proved; part (iv) is quite similar.

(v) This part follows immediately from parts (i), (ii)(f), (iii), and (iv)(f). ☐

Proof of Proposition 3.9. According to parts (ii)(a) and (iv)(a) of Proposition 3.8, event $\{\hat{x}(X,U) \neq X\}$ is contained in event $\{X \neq 0,\ \tilde{G}(X,U) \in D\}$, where D stands for the set of all points in $\mathbb{R}$ at which at least one of the monotonic functions x_+ or x_- is discontinuous. Since the set D is at most countable, $\mathsf{P}\left(\hat{x}(X,U) \neq X\right) \leqslant \sum_{h \in D} \mathsf{P}\left(X \neq 0,\ \tilde{G}(X,U) = h\right) = 0$, by Proposition 3.4. ☐

Proof of Proposition 3.10. Implications (i)$\Rightarrow$ (ii)$\Leftrightarrow$ (iii)$\Rightarrow$ (iv) follow straight from the corresponding definitions. Implication (iv)$\Rightarrow$ (v) follows by Proposition 3.9. Implication (ii)$\Rightarrow$ (i) follows by the identity

$$\mathsf{P}(X \in A) = \int_A \frac{1}{x}\, \mathrm{d}G(x)$$

for all $A \in \mathscr{B}(\mathbb{R} \setminus \{0\})$, which in turn follows from definition (2.1). It remains to prove implication (v)$\Rightarrow$ (iii). Toward this end, assume (v) and observe the equivalence

$$x_-\big(\tilde{G}(x,u)\big) \neq -x_+\big(\tilde{G}(x,u)\big) \iff \hat{x}(x,u) \neq x$$

for all $(x,u) \in \mathbb{R} \times [0,1]$ such that $\mathsf{r}(x,u) = -x$. Therefore and by Proposition 3.5, (3.12), and Proposition 3.9,

$$\begin{aligned}\frac{1}{m}\int_0^m \mathrm{I}\{x_-(h) \neq -x_+(h)\}\,\mathrm{d}h &= \mathsf{P}\left(x_-\big(\tilde{G}(Y,U)\big) \neq -x_+\big(\tilde{G}(Y,U)\big)\right) \\ &= \mathsf{P}\left(\hat{x}(Y,U) \neq X\right) = \frac{1}{2m}\,\mathsf{E}\,|X|\,\mathrm{I}\{\hat{x}(X,U) \neq X\} = 0,\end{aligned}$$

so that $x_- = -x_+$ almost everywhere on $[0,m]$ (with respect to the Lebesgue measure) and hence on an everywhere dense subset of $[0,m]$. Now it remains to recall property (iv) in Proposition 3.1, taking also into account that $x_\pm(0) = 0$. □

Proof of Proposition 3.11. First of all, v is Borel by part (i) of Remark 2.1. Next, in the case $x \geqslant 0$, one has $y = \mathsf{r}(x,u) = x_-(h)$ and, by (3.6), $G(y+) \leqslant h \leqslant G(y)$; so, $v \in [0,1]$ and $h = \tilde{G}(y,v)$, whence $\mathsf{r}\big(\mathsf{r}(x,u),v\big) = \mathsf{r}(y,v) = x_+(h) = \hat{x}(x,u)$; that is, (3.15) follows in the case $x \geqslant 0$; the case $x \leqslant 0$ is quite similar. □

Proof of Proposition 3.12. This follows immediately from Propositions 3.11 and 3.9, on letting $V := \mathsf{v}(X,U)$. □

Proof of Proposition 3.13. By monotone convergence, without loss of generality (w.l.o.g.) let us assume that the function g is bounded. Now, in view of (2.8) and the independence of X and U, observe that the difference between the left-hand side and the right-hand side of (2.12) equals $\mathsf{E}X\psi\big(X,\mathsf{r}(X,U)\big)$, where

$$\psi(x,y) := \frac{g(x,r;x,r) - g(r,x;r,x)}{x-r}\,\mathrm{I}\{xr \leqslant 0, x \neq r\}$$

for all real x and r, so that $\psi(x,r)$ is understood as 0 if $x = r$. The function ψ is symmetric, and the expression $|x\psi(x,r)| \leqslant |g(x,r;x,r) - g(r,x;r,x)|$ is bounded over all real x and r. It remains to refer to Proposition 3.15, proved later in this paper. □

Proof of Proposition 3.14. This follows immediately from Proposition 3.13 and (3.17). □

To prove Proposition 3.15, we shall use some notation and two lemmas, as follows.

For all real a and b, let

$$e_1(a,b) := e_{1,X}(a,b) := \mathsf{E}X\,\mathrm{I}\{X < a,\ \mathsf{r}(X,U) > b\} \quad \text{and} \tag{4.1}$$

$$e_2(a,b) := e_{2,X}(a,b) := \mathsf{E}X\,\mathrm{I}\{\mathsf{r}(X,U) < a,\ X > b\}. \tag{4.2}$$

Lemma 4.1. *For all real a and b such that $a \leqslant 0 \leqslant b$,*

$$e_1(a,b) = -m + G(a) \vee G(b).$$

Proof. Let us consider the following two cases.

Case 1: $G(a) > G(b)$. Then for all $x \in \mathbb{R}$ and $u \in (0,1)$

$$\begin{aligned} x < a &\implies x < 0 \;\&\; \tilde{G}(x,u) \geqslant G(x+) \geqslant G(a) > G(b) \\ &\implies x < 0 \;\&\; x_+(\tilde{G}(x,u)) > b \qquad (4.3) \\ &\implies \mathsf{r}(x,u) > b, \end{aligned}$$

where implication (4.3) follows from (3.1). So, in this case

$$\begin{aligned} e_1(a,b) = \mathsf{E}X\,\mathrm{I}\{X < a\} &= \mathsf{E}X\,\mathrm{I}\{X < 0\} - \mathsf{E}X\,\mathrm{I}\{X \in [a,0)\} \\ &= -m + G(a) = -m + G(a) \vee G(b). \end{aligned}$$

Case 2: $G(a) \leqslant G(b)$. Then for all $x \in \mathbb{R}$ and $u \in (0,1)$

$$\begin{aligned} x < 0 \;\&\; \mathsf{r}(x,u) > b &\iff x < 0 \;\&\; x_+(\tilde{G}(x,u)) > b \\ &\iff x < 0 \;\&\; \tilde{G}(x,u)) > G(b) \qquad (4.4) \\ &\implies x < 0 \;\&\; G(x) \geqslant \tilde{G}(x,u) > G(b) \geqslant G(a) \\ &\implies x < a, \end{aligned}$$

where equivalence (4.4) follows from (3.1). Also, $x < a$ implies $x < 0$, since $a \leqslant 0$. So, in Case 2 event $\{X < a,\ \mathsf{r}(X,U) > b\}$ coincides with $\{X < 0,\ \mathsf{r}(X,U) > b\}$ and hence, by (4.4), with $\{X < 0,\ \tilde{G}(X,U) > G(b)\}$. So,

$$\begin{aligned} e_1(a,b) &= \mathsf{E}X\,\mathrm{I}\{X < 0,\ \tilde{G}(X,U) > G(b)\} \\ &= \mathsf{E}X\,\mathrm{I}\{X < 0\} - \mathsf{E}X\,\mathrm{I}\{X < 0,\ \tilde{G}(X,U) \leqslant G(b)\} \\ &= -m + G(b) = -m + G(a) \vee G(b), \end{aligned}$$

where the third equality follows by (3.11). □

Lemma 4.2. *For all real a and b,*

$$\mathsf{E}X\,\mathrm{I}\{X < a, \mathsf{r}(X,U) > b\} + \mathsf{E}X\,\mathrm{I}\{\mathsf{r}(X,U) < a, X > b\} = 0.$$

Proof. We have to prove that $e_1 + e_2 = 0$ on $\mathbb{R}^2$, where e_1 and e_2 are given by (4.1) and (4.2). Observe that

$$\begin{aligned} G_{-X}(x) &= G_X(-x); \qquad (4.5) \\ \mathsf{r}_{-X}(x,u) &= -\mathsf{r}_X(-x,u); \\ e_{1,-X}(x,y) &= -e_{2,X}(-y,-x) \qquad (4.6) \end{aligned}$$

for all real x and y and $u \in (0,1)$. Let us now consider the four possible cases.

Case 1: $a \leqslant 0 \leqslant b$. Then, by (4.6), Lemma 4.1, and (4.5),

$$\begin{aligned} e_2(a,b) = e_{2,X}(a,b) = -e_{1,-X}(-b,-a) &= m - G_{-X}(-b) \vee G_{-X}(-a) \\ &= m - G_X(b) \vee G_X(a) = -e_1(a,b), \end{aligned}$$

again by Lemma 4.1. So, $e_1(a,b)+e_2(a,b)=0$ in Case 1.

Case 2: $a>0$ and $b\geqslant 0$. Then

$$e_1(a,b)=e_1(0,b), \tag{4.7}$$

since the inequalities $\mathsf{r}(X,U)>b$ and $b\geqslant 0$ imply that $\mathsf{r}(X,U)>0$ and hence $X<0$, so that $X<a$. Next, in view of condition $b\geqslant 0$ and Proposition 3.7, one has $\mathsf{P}\big(X>b,\ \mathsf{r}(X,U)=0\big)=0$. This implies $e_2(0+,b)=e_2(0,b)$, whence $e_2(a,b)=e_2(0+,b)=e_2(0,b)=-e_1(0,b)$; the third equality here follows by Case 1. Now, in view of (4.7), one concludes that $e_1(a,b)+e_2(a,b)=0$ in Case 2 as well.

Case 3: $a\leqslant 0$ and $b<0$. This case follows from Case 2 by (4.6).

Case 4: $b<0<a$. In this case, taking into account the inequality $X\,\mathsf{r}(X,U)\leqslant 0$, one has

$$e_1(a,b)=e_2(b+,a-)-e_2(0+,a-)-e_2(b+,0-),$$
$$e_2(a,b)=e_1(b+,a-)-e_1(0+,a-)-e_1(b+,0-).$$

By adding these two equalities, one obtains Case 4 from the already considered Cases 1, 2, 3. □

Proof of Proposition 3.15. We have to show that $T(\psi)=0$, where

$$T(\psi):=\mathsf{E}\,X\psi\big(X,\mathsf{r}(X,U)\big).$$

In view of the identity $\psi=\max(0,\psi)-\max(0,-\psi)$, let us assume w.l.o.g. that $\psi\geqslant 0$ on $\mathbb{R}$. Then, by the symmetry of ψ, one has the identity $\psi(x,y)=\frac12\int_0^\infty\psi_{A_t}(x,y)\,\mathrm{d}t$ for all real x and y, where $A_t:=\{(x,y)\in\mathbb{R}^2\colon \psi(x,y)\geqslant t\}$ and

$$\psi_A(x,y):=\mathrm{I}\{(x,y)\in A\}+\mathrm{I}\{(y,x)\in A\}$$

for all real x and y and all $A\in\mathscr{B}(\mathbb{R}^2)$. Hence, by Fubini's theorem, it is enough to show that the finite signed measure τ defined by the formula $\tau(A):=T(\psi_A)$ for $A\in\mathscr{B}(\mathbb{R}^2)$ is zero. So, it is enough to show that $\tau(A)=0$ for the sets A of the form $(-\infty,a)\times(b,\infty)$, for all real a and b, since the set of all such sets generates the entire σ-algebra $\mathscr{B}(\mathbb{R}^2)$. Now it remains to refer to Lemma 4.2. □

Proof of Proposition 3.16. From Theorem 2.2 and (3.23), it follows that (i) $\mathsf{E}\,\frac{\mathsf{r}(X,U)}{X}\leqslant -1$ always and (ii) $\mathsf{E}\,\frac{\mathsf{r}(X,U)}{X}=-1$ iff $\mathsf{r}(X,U)+X=0$ a.s. It remains to use the equivalence (v)$\Leftrightarrow$(i) of Proposition 3.10. □

Proof of Proposition 3.17. Let $\psi\colon\mathbb{R}^2\to\mathbb{R}$ be any nonnegative Borel function. By (3.13),

$$\mathsf{E}\,\psi\big(Y_\pm,\mathsf{r}(Y_\pm,U)\big)=\pm\frac{1}{m}\,\mathsf{E}\,X^\pm\psi\big(X,\mathsf{r}(X,U)\big)\quad\text{and}$$
$$\mathsf{E}\,\psi\big(Y,\mathsf{r}(Y,U)\big)=\frac{1}{2m}\,\mathsf{E}\,|X|E\psi\big(X,\mathsf{r}(X,U)\big)$$
$$=\tfrac12\,\mathsf{E}\,\psi\big(Y_+,\mathsf{r}(Y_+,U)\big)+\tfrac12\,\mathsf{E}\,\psi\big(Y_-,\mathsf{r}(Y_-,U)\big). \tag{4.8}$$

Letting now $g_1(x,r;\tilde x,\tilde r)\equiv|x|\psi(x,r)$ and $g_2(x,r;\tilde x,\tilde r)\equiv|x|\psi(r,x)$ in Proposition 3.14, one has $\mathsf{E}\,\psi\big(Y,\mathsf{r}(Y,U)\big)=\mathsf{E}\,\psi\big(\mathsf{r}(Y,U),Y\big)$, which proves (3.24).

The first equality in (3.25) is proved similarly, with $g_1(x,r;\tilde{x},\tilde{r}) \equiv x^+\psi(x,r)$ and $g_2(x,r;\tilde{x},\tilde{r}) \equiv -x^-\psi(r,x)$.

To prove the second equality in (3.25), let $H := \tilde{G}(Y_-,U)$; then, by Proposition 3.5, the r.v. H is indeed uniformly distributed in $[0,m]$. Recall also that $Y_- \leqslant 0$ a.s. Therefore and in view of (3.14) and Proposition 3.9, $x_-(H) = x_-\big(\tilde{G}(Y_-,U)\big) = \hat{x}(Y_-,U) \overset{\text{a.s.}}{=} Y_-$. On the other hand, by (2.6), $x_+(H) = x_+\big(\tilde{G}(Y_-,U)\big) = \mathsf{r}(Y_-,U)$. Hence, $\big(\mathsf{r}(Y_-,U),Y_-\big) \overset{\text{a.s.}}{=} \big(x_+(H),x_-(H)\big)$.

The second and third equalities in (3.26) follow immediately from (3.25). In turn, these two equalities imply the first equality in (3.26), in view of (4.8) (used with symmetric ψ).

The rest of Proposition 3.17 follows immediately from (3.24) and (3.25), except for the "except when" statement in the parentheses. To prove this latter statement, note that, in view of the inequality $\frac{a}{b}+\frac{b}{a} < -2$ for all real a and b with $ab<0$ and $a \neq -b$, the equality $\mathsf{E}\,\frac{\mathsf{r}(Y,U)}{Y} = \mathsf{E}\,\frac{Y}{\mathsf{r}(Y,U)}$ implies that $\mathsf{E}\,\frac{\mathsf{r}(Y,U)}{Y} = \mathsf{E}\,\frac{Y}{\mathsf{r}(Y,U)} = \frac{1}{2}\big(\mathsf{E}\,\frac{\mathsf{r}(Y,U)}{Y} + \mathsf{E}\,\frac{Y}{\mathsf{r}(Y,U)}\big) < -1$ unless $\mathsf{r}(Y,U) = -Y$ a.s. It remains now to refer to (3.12) and Proposition 3.10. □

Proof 1 of Proposition 3.18. W.l.o.g. the function g is bounded (by monotone convergence) and nonnegative $\big($by the identity $g = \max(0,g) - \max(0,-g)\big)$. Write $g(x) = \int_0^\infty g_{A_t}(x)\,\mathrm{d}t$ for all $x \in \mathbb{R}$, where $g_A(x) := \mathrm{I}\{x \in A\}$ and $A_t := \{x \in \mathbb{R}\colon g(x) \geqslant t\}$. So, by Fubini's theorem, w.l.o.g. $g = g_A$ for some $A \in \mathscr{B}(\mathbb{R}\setminus\{0\})$. Let then $\lambda(A)$ and $\rho(A)$ denote, respectively, the left-hand side $\big($say $L(g)\big)$ and the right-hand side $\big($say $R(g)\big)$ of (3.27) with $g = g_A$. It remains to show that the measures λ and ρ coincide on $\mathscr{B}(\mathbb{R}\setminus\{0\})$. Since the sets A of the form $(-\infty,-b)$ or (b,∞) for some $b>0$ generate the σ-algebra $\mathscr{B}(\mathbb{R}\setminus\{0\})$, it suffices to show that $L\big(g_{(-\infty,-b)}\big) = R\big(g_{(-\infty,-b)}\big)$ and $L\big(g_{(b,\infty)}\big) = R\big(g_{(b,\infty)}\big)$ for all $b>0$.

Let next $g_a(x) := x\,\mathrm{I}\{0 < x \leqslant a\}$ and observe that

$$\int_0^\infty g_a(x)\,\nu(\mathrm{d}a) = x^+\,\nu\big([x,\infty)\big) = g_{(b,\infty)}(x) \tag{4.9}$$

for all $x \in \mathbb{R}$ if $\nu = \nu_b$, where ν_b is the finite signed measure on $(0,\infty)$ uniquely determined by the condition that $x\,\nu_b\big([x,\infty)\big) = \mathrm{I}\{x > b\}$ for all $x \in \mathbb{R}$.

On the other hand, for any $a>0$, one has $L(g_a) = \mathsf{E}\,g_a(X) = G(a)$ $\big($by (2.1)$\big)$ and

$$R(g_a) = \int_0^m \mathrm{I}\{x_+(h) \leqslant a\}\,\mathrm{d}h = \int_0^m \mathrm{I}\{G(a) \geqslant h\}\,\mathrm{d}h = G(a)$$

$\big($by (3.1)$\big)$. So, $L(g_a) = R(g_a)$ for all $a>0$.

Observe also that $\int_0^\infty |g_a(x)|\,|\nu_b(\mathrm{d}a)| \leqslant c_b\,x^+$ for all $x \in \mathbb{R}$ and $b>0$, where $c_b := \int_0^\infty |\nu_b(\mathrm{d}a)| < \infty$. So, again by Fubini's theorem $\big($and in view of (4.9)$\big)$, it follows that $L(g_{(b,\infty)}) = R(g_{(b,\infty)})$ for all $b>0$. Similarly, $L(g_{(-\infty,-b)}) = R(g_{(-\infty,-b)})$ for all $b>0$. □

Proof 2 of Proposition 3.18. W.l.o.g. the function g in Proposition 3.18 is nonnegative (otherwise, consider its positive and negative parts). Let $\psi(x) := g(x)/|x|$ for all real $x \neq 0$ and $\psi(0) := 0$.

Then, by (3.13) and [3, Theorem 2.2],

$$\begin{aligned} \mathsf{E}\, g(X) = 2m\, \mathsf{E}\, \psi(Y) = 2m \int_0^1 \Big(\tfrac{1}{2}\, \psi\big(Y_1(t)\big) + \tfrac{1}{2}\, \psi\big(Y_2(t)\big)\Big)\, \mathrm{d}t \\ = m \int_0^1 \Big(\frac{g\big(Y_1(t)\big)}{|Y_1(t)|} + \frac{g\big(Y_2(t)\big)}{|Y_2(t)|}\Big)\, \mathrm{d}t = m \int_0^1 \mathsf{E}\, g(Z_t)\, \frac{\mathrm{d}t}{\mathsf{E}\, Z_t^+}, \end{aligned} \tag{4.10}$$

where $Z_t := X_{Y_1(t),Y_2(t)}$, $Y_1(t) := F^{-1}(\frac{t}{2})$, $Y_2(t) := F^{-1}(1-\frac{t}{2})$, $F^{-1}(u) := \inf\{y \in \mathbb{R}\colon F(y) \geqslant u\}$, and, with $G = G_X$,

$$F(y) := \mathsf{P}(Y \leqslant y) = \begin{cases} \frac{1}{2} - \frac{1}{2m}\, G(y+) & \text{if } y \leqslant 0, \\ \frac{1}{2} + \frac{1}{2m}\, G(y) & \text{if } y \geqslant 0, \end{cases}$$

the latter equality taking place in view of (3.12) and (2.1).

Next, fix any $t \in (0,1)$. Then $1 - \frac{t}{2} > \frac{1}{2}$, and so,

$$\begin{aligned} Y_2(t) = \inf\{y \geqslant 0\colon \tfrac{1}{2} + \tfrac{1}{2m}\, G(y) \geqslant 1 - \tfrac{t}{2}\} \\ = \inf\{y \geqslant 0\colon G(y) \geqslant m(1-t)\} = x_+\big(m(1-t)\big). \end{aligned} \tag{4.11}$$

Also, $\frac{t}{2} < \frac{1}{2}$, whence, letting $h := m(1-t)$, one has $h \in (0,m)$ and

$$\begin{aligned} Y_1(t) = \inf\{y \leqslant 0\colon \tfrac{1}{2} - \tfrac{1}{2m}\, G(y+) \geqslant \tfrac{t}{2}\} \\ = \inf\{y \leqslant 0\colon G(y+) \leqslant h\}. \end{aligned}$$

Therefore, $Y_1(t) \in (-\infty, 0]$ (since $G(y+) \underset{y\to-\infty}{\longrightarrow} m > h$), $G\big(Y_1(t)+\big) \leqslant h$ $\big($since the function $x \mapsto G(x+)$ is right-continuous on $[-\infty,0]\big)$, $G(y+) > h$ for all $y < Y_1(t)$, and so, $G\big(Y_1(t)\big) \geqslant h$. Now (3.2) yields $Y_1(t) \leqslant x_-(h)$. If at that $Y_1(t) < x_-(h)$ then $G\big(Y_1(t)+\big) \geqslant G\big(x_-(h)\big) \geqslant h \geqslant G\big(Y_1(t)+\big)$, which implies that $\tilde{G}\big(Y_1(t),0\big) = G\big(Y_1(t)+\big) = h$; therefore, by (3.14), $\hat{x}\big(Y_1(t),0\big) = x_-\Big(\tilde{G}\big(Y_1(t),0\big)\Big) = x_-(h) > Y_1(t)$. Hence, by part (iv)(a) of Proposition 3.8, $Y_1(t) = x_-(h)$ unless $h = m(1-t)$ is a point of discontinuity of the nonincreasing function x_-. Thus, $Y_1(t) = x_-\big(m(1-t)\big)$ for almost all $t \in (0,1)$. Now (3.27) follows in view of (4.10) and (4.11). □

Proof of Proposition 3.19. This is quite similar to the proof of Proposition 3.5. □

Proof of Proposition 3.20. Let $g\colon \mathbb{R} \to \mathbb{R}$ is any Borel function bounded from below (or from above). In addition to the function Ψ_g defined by (3.28), introduce the functions $\Psi_{g,+}$ and $\Psi_{g,-}$ defined by the formulas

$$\Psi_{g,+}(h) := \frac{\mathsf{E}\, g(X_h)}{x_+(h)} \quad \text{and} \quad \Psi_{g,+}(h) := -\frac{\mathsf{E}\, g(X_h)}{x_-(h)}$$

for all $h \in (0,m)$, so that

$$\Psi_g = \Psi_{g,+} + \Psi_{g,-}.$$

Consider now the case $g(0) = 0$. Then, in view of (3.29), Proposition 3.5, (3.13), (2.6), Proposition 3.9, and (3.17),

$$
\begin{aligned}
\mathsf{E}\, g(X) &= m\,\mathsf{E}\,\Psi_{g,+}(H) + m\,\mathsf{E}\,\Psi_{g,-}(H) \\
&= m\,\mathsf{E}\,\Psi_{g,+}\big(\tilde{G}(Y_+,U)\big) + m\,\mathsf{E}\,\Psi_{g,-}\big(\tilde{G}(Y_-,U)\big) \\
&= \mathsf{E} X^+ \Psi_{g,+}\big(\tilde{G}(X,U)\big) + \mathsf{E}(-X^-)\Psi_{g,-}\big(\tilde{G}(X,U)\big) \\
&= \int_{(0,\infty)\times[0,1]} \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u \\
&\qquad \times \frac{g\Big(x_-\big(\tilde{G}(x,u)\big)\Big)\, x_+\big(\tilde{G}(x,u)\big) - g\Big(x_+\big(\tilde{G}(x,u)\big)\Big)\, x_-\big(\tilde{G}(x,u)\big)}{x_+\big(\tilde{G}(x,u)\big) - x_-\big(\tilde{G}(x,u)\big)} \\
&\quad + \int_{(-\infty,0)\times[0,1]} \mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u \\
&\qquad \times \frac{g\Big(x_+\big(\tilde{G}(x,u)\big)\Big)\, x_-\big(\tilde{G}(x,u)\big) - g\Big(x_-\big(\tilde{G}(x,u)\big)\Big)\, x_+\big(\tilde{G}(x,u)\big)}{x_-\big(\tilde{G}(x,u)\big) - x_+\big(\tilde{G}(x,u)\big)} \\
&= \Big(\int_{(0,\infty)\times[0,1]} + \int_{(-\infty,0)\times[0,1]}\Big) \frac{g\big(\mathrm{r}(x,u)\big)\,x - g(x)\,\mathrm{r}(x,u)}{x-\mathrm{r}(x,u)}\,\mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u \\
&= \int_{\mathbb{R}\times[0,1]} \mathsf{E}\, g\big(X_{x,\mathrm{r}(x,u)}\big)\,\mathsf{P}(X\in \mathrm{d}x)\,\mathrm{d}u.
\end{aligned}
\tag{4.12}
$$

So, identity (3.30) is proved in the case when $g(0) = 0$. But for $g(x) \equiv \mathrm{I}\{x = 0\}$, (3.30) follows by Proposition 3.7. So, (3.30) is completely proved. The proofs of (3.31) and (3.32) are similar, but using $m\,\mathsf{E}\,\Psi_{g,+}\big(\tilde{G}(Y_-,U)\big) + m\,\mathsf{E}\,\Psi_{g,-}\big(\tilde{G}(Y_+,U)\big)$ and $2m\,\mathsf{E}\,\Psi_g\big(\tilde{G}(Y,U)\big)$ instead of $m\,\mathsf{E}\,\Psi_{g,+}\big(\tilde{G}(Y_+,U)\big) + m\,\mathsf{E}\,\Psi_{g,-}\big(\tilde{G}(Y_-,U)\big)$ in (4.12); (3.32) is also an obvious corollary of (3.30) and (3.31). $\square$

Proof of Proposition 3.21. Since k is bounded from below, w.l.o.g. one has $k \geqslant 0$. Then w.l.o.g. $\mathsf{E}\,k(\tilde{X}_1,\tilde{X}_2) < \infty$, since otherwise inequality (3.35) is trivial. Just to simplify writing, assume that $I_1 = I_2 = [0,\infty)$. Then $0 \leqslant k(\tilde{X}_1,0) + k(0,\tilde{X}_2) \leqslant k(\tilde{X}_1,\tilde{X}_2) + k(0,0)$ a.s. (by the superadditivity), whence the r.v.'s $k(\tilde{X}_1,0)$ and $k(0,\tilde{X}_2)$ are integrable, and so are $k(X_1,0)$ and $k(0,X_2)$, by (3.34); moreover, $\mathsf{E}\,k(X_1,0) = \mathsf{E}\,k(\tilde{X}_1,0)$ and $\mathsf{E}\,k(0,X_2) = \mathsf{E}\,k(0,\tilde{X}_2)$.

Let μ_k be the nonnegative measure on $\mathscr{B}\big((0,\infty)^2\big)$ defined by the formula

$$\mu_k\big((a,b]\times(c,d]\big) := k(a,c) + k(b,d) - k(a,d) - k(b,c)$$

for all a,b,c,d in $[0,\infty)$ such that $a < b$ and $c < d$. Then

$$k(X_1,X_2) = k(X_1,0) + k(0,X_2) - k(0,0) + \iint_{(0,\infty)^2} \mathrm{I}\{x_1 \leqslant X_1, x_2 \leqslant X_2\}\,\mu_k(\mathrm{d}x_1,\mathrm{d}x_2) \tag{4.13}$$

a.s. Hence, by Fubini's theorem,

$$
\begin{aligned}
\mathsf{E}\,k(X_1,X_2) &= \mathsf{E}\,k(X_1,0) + \mathsf{E}\,k(0,X_2) - k(0,0) \\
&\qquad + \iint_{(0,\infty)^2} \mathsf{E}\,\mathrm{I}\{x_1 \leqslant X_1, x_2 \leqslant X_2\}\,\mu_k(\mathrm{d}x_1,\mathrm{d}x_2).
\end{aligned}
$$

A similar equality holds with $\tilde{X}_1$ and $\tilde{X}_2$ in place of X_1 and X_2. Recall that $\mathsf{E}\,k(X_1,0) = \mathsf{E}\,k(\tilde{X}_1,0)$ and $\mathsf{E}\,k(0,X_2) = \mathsf{E}\,k(0,\tilde{X}_2)$. It remains to observe that $\mathsf{E}\,\mathrm{I}\{x_1 \leqslant X_1, x_2 \leqslant X_2\} \leqslant \mathsf{P}(x_1 \leqslant X_1) \wedge \mathsf{P}(x_2 \leqslant X_2) = \mathsf{P}(x_1 \leqslant \tilde{X}_1) \wedge \mathsf{P}(x_2 \leqslant \tilde{X}_2) = \mathsf{E}\,\mathrm{I}\{x_1 \leqslant \tilde{X}_1, x_2 \leqslant \tilde{X}_2\}$ for all x_1 and x_2, where the latter equality is easy to deduce from (3.33); alternatively, inequality $\mathsf{E}\,\mathrm{I}\{x_1 \leqslant X_1, x_2 \leqslant X_2\} \leqslant \mathsf{E}\,\mathrm{I}\{x_1 \leqslant \tilde{X}_1, x_2 \leqslant \tilde{X}_2\}$ follows (say) by [33, Theorem 2], since $(z_1,z_2) \mapsto \mathrm{I}\{x_1 \leqslant z_1, x_2 \leqslant z_2\}$ is a *bounded* right-continuous superadditive function. □

Proof of Proposition 3.22. The proof is quite similar to that of [33, Corollary 2.2(a)]. We shall only indicate the necessary changes in that proof, in the notations used there, including the correction of a couple of typos: use the interval $I_\varepsilon := (\varepsilon, \frac{1}{\varepsilon}]$ with $\varepsilon \downarrow 0$ instead of $(-B,B]$ and, accordingly, replace QB by I_ε^2; w.l.o.g. one may assume here that $h = 0$; one does not need to assume that the integral $\int \varphi\, dH$ at the end of [33, page 819] is finite; on line 1 of [33, page 820], there should be $\int (h-\varphi)\, d\overline{H}$ and $\liminf$ instead of $\int (h-\varphi)\, dH$ and $\lim$, respectively. □

Proof of Proposition 3.23.

(i) In view of (3.13), (3.36) $\big($with $X^+ g(X)$ in place of $g(X)\big)$, (2.8), and (3.37),

$$\mathsf{E}\,g(Y_+) = \frac{1}{m}\,\mathsf{E} X^+\, g(X) = \frac{1}{m}\int_S \frac{y_+(s)\, g\big(y_+(s)\big)\, y_-(s)}{y_-(s) - y_+(s)}\,\nu(\mathrm{d}s) = \int_S g\big(y_+(s)\big)\,\tilde{\nu}(\mathrm{d}s) \tag{4.14}$$

for any bounded Borel function $g\colon \mathbb{R} \to \mathbb{R}$. In particular, letting here $g \equiv 1$, one sees that $\tilde{\nu}$ is a probability measure, which proves part (i) of the proposition.

(ii) Identity (4.14) means that $Y_+ \overset{\mathrm{D}}{=} y_+$. Similarly, $Y_- \overset{\mathrm{D}}{=} y_-$. This proves part (ii) of the proposition. It remains to prove part

(iii) The inequality in (3.38) follows immediately from Propositions 3.21, 3.22, and the just proved part (ii) of Proposition 3.23. The equalities in (3.38) follow immediately from relations (3.25) and (3.26) in Proposition 3.17.

As explained in Remark 3.24, relations (3.39), (3.41), and (3.42) are special cases of (3.38). Next, (3.40) follows immediately from (3.39), (3.26), and (3.24). Finally, (3.43) follows from (3.38) in view of Proposition 3.21 with $I_1 = I_2 = [0,\infty)$ by taking $k(y_1,y_2) \equiv (y_1 + y_2 + \varepsilon)^p$ for $\varepsilon > 0$, and then letting $\varepsilon \downarrow 0$ and using the monotone convergence theorem.

So, part (iii) and thus the entire Proposition 3.23 are proved. □

Proof of Proposition 3.25. The first identity in (3.44) is a special case of Proposition 3.18, with $g(x) \equiv \mathrm{I}\{x > 0\}$; the second identity is quite similar. □

Proof of Proposition 3.26. Let

$$L(x) := \begin{cases} \sup\{h \in [0,m]\colon y_+(h) \leqslant x\} & \text{if } x \in [0,\infty], \\ \sup\{h \in [0,m]\colon -y_-(h) \leqslant -x\} & \text{if } x \in [-\infty,0]. \end{cases} \tag{4.15}$$

Then one can check that relations (2.2), (3.1), and (3.2) hold with the functions L and $y_\pm$ in place of G and $x_\pm$, respectively. Introduce also a nonnegative measure ν on $\mathscr{B}(\mathbb{R})$ by the formula

$$\nu(A) := \int_{A\setminus\{0\}} \left|\tfrac{1}{x}\,\mathrm{d}L(x)\right| \tag{4.16}$$

for all $A \in \mathscr{B}(\mathbb{R})$, so that $\big($cf. (2.1)$\big)$

$$L(x) = \begin{cases} \int_{(0,x]} z\,\nu(\mathrm{d}z) & \text{if } x \in [0,\infty], \\ \int_{[x,0)} (-z)\,\nu(\mathrm{d}z) & \text{if } x \in [-\infty,0]. \end{cases} \tag{4.17}$$

Then, using (3.1) (with L and y_+ instead of G and x_+) and Fubini's theorem, one has

$$\int_0^m \frac{\mathrm{d}h}{y_+(h)} = \int_0^m \mathrm{d}h \int_0^\infty \mathrm{I}\{y_+(h) \leqslant \tfrac{1}{t}\}\,\mathrm{d}t = \int_0^m \mathrm{d}h \int_0^\infty \mathrm{I}\{L(\tfrac{1}{t}) \geqslant h\}\,\mathrm{d}t =$$
$$\int_0^\infty L(\tfrac{1}{t})\,\mathrm{d}t = \int_0^\infty \mathrm{d}t \int_{\mathbb{R}} x\,\mathrm{I}\{0 < x \leqslant \tfrac{1}{t}\}\,\nu(\mathrm{d}x) = \int_{\mathbb{R}} \mathrm{I}\{x > 0\}\,\nu(\mathrm{d}x) = \nu\big((0,\infty)\big).$$

Similarly, $\int_0^m \dfrac{\mathrm{d}h}{-y_-(h)} = \nu\big((-\infty,0)\big)$. So, by condition(3.45), one has $\nu(\mathbb{R}\setminus\{0\}) \leqslant 1$. So, there exists a unique probability distribution μ on $\mathbb{R}$ such that $\mu(A) = \nu(A)$ for all $A \in \mathscr{B}(\mathbb{R}\setminus\{0\})$. Let X be any r.v. with this distribution μ. Then, by (4.17) and (4.15), one has $\mathsf{E}X^+ = L(\infty) = m = L(-\infty) = \mathsf{E}(-X^-)$, whence $\mathsf{E}X = 0$. Also, $G_\mu = L$, and so, in view of (3.1) and (3.2), the functions $x_\pm$ for the zero-mean distribution μ coincide with $y_\pm$. The uniqueness of μ follows because (i) the functions $x_\pm$ uniquely determine the function G $\big($via (3.1) and (3.2)$\big)$ and (ii) the function G uniquely determines the distribution $\big($cf. (4.16)$\big)$. □

Proof of Proposition 3.27.

$\boxed{\text{Checking (I)} \Longrightarrow \text{(II).}}$ Here it is assumed that there exists a zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$ whose reciprocating function $\mathsf{r} := \mathsf{r}_\mu$ satisfies conditions $\mu_+ = \nu$ and $\mathsf{r}_+ = \mathsf{s}$. We have to show at this point that then conditions (a)–(k) necessarily take place.

(a) Since $\mathsf{r}_+ = \mathsf{s}$, one has $\mathsf{s}(0,u) = \mathsf{r}(0,u) = 0$ for all $u \in [0,1]$, by definition (2.6).

(b) The conditions $\mathsf{r} = \mathsf{r}_\mu$, $\mu_+ = \nu$, and $\mathsf{r}_+ = \mathsf{s}$ imply that

$$\mathsf{s}(x,u) = x_{-,\mu}\big(\tilde{G}_\mu(x,u)\big) \quad \text{for all} \quad (x,u) \in [0,\infty] \times [0,1]. \tag{4.18}$$

So, condition (b) follows by Proposition 3.2, since $x_{-,\mu}$ is nonincreasing on $[0,m_\mu]$ $\big($by part (i) of Proposition 3.1$\big)$.

(c), (d) These conditions follow by (4.18), Proposition 3.2, and property (iv) in Proposition 3.1, because $\tilde{G}_\mu(x,0) = G_\mu(x-)$ is left-continuous in $x \in [0,\infty]$ and $\tilde{G}_\mu(x,u)$ is affine and hence continuous in $u \in [0,1]$ for every $x \in [0,\infty]$. $\big($Note that, if $\tilde{G}(x,u) = 0$ for some $(x,u) \in (0,\infty] \times [0,1]$, then $\tilde{G}(z,v) = 0$ for all (z,v) such that $(0,0) \prec (z,v) \prec (x,u)$.$\big)$

(e), (f) The necessity of these conditions is obvious.

(g) If $\nu(\{x\}) = 0$ and $x \in (0,\infty]$ then $\mu(\{x\}) = 0$, $\tilde{G}_\mu(x,1) = \tilde{G}_\mu(x,0)$, and so, by (4.18), $\mathsf{s}(x,1) = \mathsf{s}(x,0)$.

(h) If $\nu\big((x,y)\big) = 0$ and $0 \leqslant x < y \leqslant \infty$ then $\mu\big((x,y)\big) = 0$, $G_\mu(y-) - G_\mu(x) = \int_{(x,y)} z\,\mu(\mathrm{d}z) = 0$, and so, by (4.18), $\mathsf{s}(x,1) = x_{-,\mu}\big(G_\mu(x)\big) = x_{-,\mu}\big(G_\mu(y-)\big) = \mathsf{s}(y,0)$.

(i) This follows from (4.18) and property (ii) in Proposition 3.1.

(j) Similarly, this follows from (4.18) and property (iii) in Proposition 3.1.

(k) By Proposition 3.5, the r.v. $\tilde{G}_\mu(Y_{+,\mu}, U)$ is uniformly distributed in the interval $[0, m_\nu]$, where $Y_{+,\mu}$ is any r.v. which is independent of U and such that (cf. (3.12)) $\mathsf{P}(Y_{+,\mu} \in A) = \frac{1}{m_\nu} \int_A x^+ \mu(\mathrm{d}x)$ for all $A \in \mathscr{B}(\mathbb{R})$. Therefore and in view of (4.18),

$$\int_0^{m_\nu} \frac{\mathrm{d}h}{x_{-,\mu}(h)} = m\,\mathsf{E}\,\frac{1}{x_{-,\mu}(\tilde{G}_\mu(Y_{+,\mu}, U))} = \int_{(0,\infty)\times[0,1]} \frac{x}{\mathsf{s}(x,u)}\,\nu(\mathrm{d}x)\,\mathrm{d}u \tag{4.19}$$

(cf. (3.13)). On the other hand, by Proposition 3.25, $\int_0^{m_\nu} \frac{\mathrm{d}h}{x_{-,\mu}(h)} = -\mu\big((-\infty, 0)\big)$. Thus,

$$\int_{(0,\infty)\times[0,1]} \Big(1 - \frac{x}{\mathsf{s}(x,u)}\Big)\,\nu(\mathrm{d}x)\,\mathrm{d}u = \nu\big((0,\infty)\big) + \mu\big((-\infty,0)\big) = \mu\big(\mathbb{R}\setminus\{0\}\big) \leqslant 1, \tag{4.20}$$

so that the necessity of condition (k) is verified.

$\boxed{\text{Checking (II)} \implies \text{(I).}}$

$\boxed{\text{Step 1.}}$ Here (assuming the conditions (a)–(k) to hold) we shall show that there exists a unique function $y_- : [0, m_\nu] \to \mathbb{R}$ such that (cf. (4.18))

$$\mathsf{s}(x,u) = y_-\big(\tilde{G}_\nu(x,u)\big) \quad \text{for all} \quad (x,u) \in [0,\infty] \times [0,1]. \tag{4.21}$$

Toward this end, let us first observe that the "range" $\tilde{G}_\nu([0,\infty]\times[0,1])$ contains the entire interval $[0, m_\nu]$. Indeed, for any given $h \in [0, m_\nu]$, let $x := x_{+,\nu}(h)$, so that $x \in [0,\infty]$. Then (cf. (3.4)) $G_\nu(x-) \leqslant h \leqslant G_\nu(x)$. Hence, $h = \tilde{G}_\nu(x,u)$ for some $u \in [0,1]$. (Here, we used (2.5) and, tacitly, condition (f).)

Now, to complete Step 1 it is enough to show for all points (x,u) and (y,v) in $[0,\infty]\times[0,1]$ one has the implication

$$\tilde{G}_\nu(x,u) = \tilde{G}_\nu(y,v) \implies \mathsf{s}(x,u) = \mathsf{s}(y,v). \tag{4.22}$$

Let us assume that indeed $\tilde{G}_\nu(x,u) = \tilde{G}_\nu(y,v)$; we have to show that $\mathsf{s}(x,u) = \mathsf{s}(y,v)$. W.l.o.g. let us also assume that here $(x,u) \prec (y,v)$, whence, by condition (b), $\mathsf{s}(x,u) \geqslant \mathsf{s}(y,v)$. One of the following two cases must take place.

Case 1: $x = y$ *and* $u < v$. Then $\tilde{G}_\nu(x,u) = \tilde{G}_\nu(y,v)$ implies, by (2.5), that $x = 0$ or $\nu(\{x\}) = 0$, whence, by conditions (a) and (g), $\mathsf{s}(x,1) = \mathsf{s}(x,0)$, so that, in view of conditions $x = y$ and (b), $\mathsf{s}(y,v) = \mathsf{s}(x,v) \geqslant \mathsf{s}(x,1) = \mathsf{s}(x,0) \geqslant \mathsf{s}(x,u)$. This, together with the inequality $\mathsf{s}(x,u) \geqslant \mathsf{s}(y,v)$, yields $\mathsf{s}(x,u) = \mathsf{s}(y,v)$.

Case 2: $x < y$. Then

$$\tilde{G}_\nu(x,u) \leqslant \tilde{G}_\nu(x,1) = G_\nu(x) \leqslant G_\nu(y-) = \tilde{G}_\nu(y,0) \leqslant \tilde{G}_\nu(y,v), \tag{4.23}$$

and so, $\tilde{G}_\nu(x,u) = \tilde{G}_\nu(y,v)$ implies that all the three inequalities in (4.23) are in fact equalities. In particular, one has $\tilde{G}_\nu(x,u) = \tilde{G}_\nu(x,1)$, which implies, by Case 1, that

$$\mathsf{s}(x,u) = \mathsf{s}(x,1). \tag{4.24}$$

Similarly, one has $\tilde{G}_\nu(y,0) = \tilde{G}_\nu(y,v)$, which implies that

$$\mathsf{s}(y,v) = \mathsf{s}(y,0). \tag{4.25}$$

To conclude Step 1, recall that one also has $G_\nu(x) = G_\nu(y-)$, whence $0 = G_\nu(y-) - G_\nu(x) = \int_{(x,y)} z\,\nu(\mathrm{d}z)$, so that $\nu\big((x,y)\big) = 0$, and, by condition (h), $\mathsf{s}(x,1) = \mathsf{s}(y,0)$. This, together with (4.24) and (4.25), yields the desired conclusion $\mathsf{s}(x,u) = \mathsf{s}(y,v)$ of implication (4.22).

$\boxed{\text{Step 2.}}$ Here (again assuming the conditions (a)–(k) to hold) we shall show that all the conditions (i)–(iv) in Proposition 3.1 are satisfied with y_- (defined by (4.21)) and m_ν in place of x_- and m.

(i) That y_- is non-increasing on $[0, m_\nu]$ follows because (by Proposition 3.2) $\tilde{G}_\nu$ is $\prec$-non-decreasing on $[0,\infty] \times [0,1]$ and, by condition (b), s is $\prec$-non-increasing. Indeed, take any h_1 and h_2 such that $0 \leqslant h_1 < h_2 \leqslant m_\nu$. Then there exist some points (x_1,u_1) and (x_2,u_2) in $[0,\infty] \times [0,1]$ such that $\tilde{G}(x_1,u_1) = h_1$ and $\tilde{G}(x_2,u_2) = h_2$; at that, necessarily $(x_1,u_1) \prec (x_2,u_2)$, and so, $y_-(h_1) = \mathsf{s}(x_1,u_1) \geqslant \mathsf{s}(x_2,u_2) = y_-(h_2)$.

(ii) That y_- is finite on $[0, m_\nu)$ follows by condition (i). Indeed, take any $h \in [0, m_\nu)$ and any $(x,u) \in [0,\infty] \times [0,1]$ such that $\tilde{G}_\nu(x,u) = h$. Then $x < \infty$, since $\tilde{G}_\nu(\infty,u) = G_\nu(\infty) = m_\nu \neq h$ for all $u \in [0,1]$. Hence, by (i), $\mathsf{s}(x,u) > -\infty$. Also, $\mathsf{s}(x,u) \leqslant 0$, in view of conditions (a) and (b). So, by (4.21), $y_-(h) = y_-\big(\tilde{G}_\nu(x,u)\big) = \mathsf{s}(x,u) \in (-\infty, 0]$.

(iii) First at this point, note that, by (4.21) and condition (a), $y_-(0) = y_-\big(\tilde{G}_\nu(0,0)\big) = \mathsf{s}(0,0) = 0$. Next, take any $h \in (0, m_\nu]$ and any $(x,u) \in [0,\infty] \times [0,1]$ such that $\tilde{G}_\nu(x,u) = h$. Then $x > 0$, since $\tilde{G}_\nu(0,u) = 0$ for all $u \in [0,1]$. Hence, again by (4.21), $y_-(h) = y_-\big(\tilde{G}_\nu(x,u)\big) = \mathsf{s}(x,u) < 0$, in view of condition (j). Thus, indeed $-y_- > 0$ on $(0, m_\nu]$.

(iv) To complete Step 2 of the proof of Proposition 3.27, we have to show that the function y_- defined by (4.21) is left-continuous on $(0, m_\nu]$. For brevity, let here $x_h := x_{+,\nu}(h)$ for all h. As was seen before, for each $h \in [0, m_\nu]$ there exists some $u_h \in [0,1]$ such that $\tilde{G}_\nu(x_h,u_h) = h$. Take any $h_0 \in (0, m_\nu]$. Then $x_{h_0} > 0$; cf. property (iii) in Proposition 3.1. One of the following two cases must take place.

Case 1: $p_0 := \nu(\{x_{h_0}\}) > 0$ and $u_{h_0} > 0$. Take any $\delta \in \big(0, \min(h_0, p_0\, x_{h_0} u_0)\big)$. Take next any $h \in (h_0 - \delta, h_0]$, so that $h \in (0, h_0]$ and (cf. property (i) in Proposition 3.1) $x_h \leqslant x_{h_0}$. In fact, we claim that $x_h = x_{h_0}$: otherwise, one would have $x_h < x_{h_0}$, whence $h = \tilde{G}_\nu(x_h,u_h) \leqslant G(x_{h_0}-) = \tilde{G}_\nu(x_{h_0},u_{h_0}) - p_0\, x_{h_0} u_0 = h_0 - p_0\, x_{h_0} u_0 < h_0 - \delta < h$, a contradiction. So, for all $h \in (h_0 - \delta, h_0]$, one has $\tilde{G}_\nu(x_{h_0},u_h) = h$ and $y_-(h) = \mathsf{s}(x_{h_0},u_h)$ (by (4.21)). Moreover, u_h is increasing in $h \in (h_0 - \delta, h_0]$, since $\tilde{G}_\nu(x_{h_0},u_h) = \tilde{G}_\nu(x_h,u_h) = h$ for all $h \in (h_0 - \delta, h_0]$. So, by condition (d), $y_-(h) = \mathsf{s}(x_{h_0},u_h) \to \mathsf{s}(x_{h_0},u_{h_0}) = y_-(h_0)$ as $h \uparrow h_0$.

Case 2: $p_0 = 0$ or $u_{h_0} = 0$. Then, with x_h and u_h defined as above, one has $h_0 = \tilde{G}_\nu(x_{h_0},u_{h_0}) = G(x_{h_0}-) = \tilde{G}_\nu(x_{h_0},0)$, so that $y_-(h_0) = \mathsf{s}(x_{h_0},0)$. Moreover, for every $h \in (0,h_0)$ one has $0 \leqslant x_h < x_{h_0}$ – because the inequality $x_{h_0} \leqslant x_h$ would imply that $h_0 = G(x_{h_0}-) \leqslant G(x_h-) = \tilde{G}_\nu(x_h,0) \leqslant \tilde{G}_\nu(x_h,u_h) = h$, which contradicts the condition $h \in (0,h_0)$. Hence and by condition (b), $\mathsf{s}(x_{h_0},0) \leqslant \mathsf{s}(x_h,u_h) \leqslant \mathsf{s}(x_h,0)$. Since $x_h = x_{+,\nu}(h)$ is left-continuous and non-decreasing in $h \in (0, m_\nu)$ (cf. properties (i) and (iv) in Proposition 3.1), it follows by condition (c) that $\mathsf{s}(x_h,0) \to \mathsf{s}(x_{h_0},0)$ as $h \uparrow h_0$. Therefore, by virtue of inequalities $\mathsf{s}(x_{h_0},0) \leqslant \mathsf{s}(x_h,u_h) \leqslant \mathsf{s}(x_h,0)$, one has $\mathsf{s}(x_h,u_h) \to \mathsf{s}(x_{h_0},0)$ as $h \uparrow h_0$. In view of equalities (4.21), $h \equiv \tilde{G}_\nu(x_h,u_h)$, and $y_-(h_0) = \mathsf{s}(x_{h_0},0)$, this means that $y_-(h) \to y_-(h_0)$ as $h \uparrow h_0$.

This completes Step 2 of the proof.

$\boxed{\text{Step 3.}}$ Now we are prepared to complete the entire proof of Proposition 3.27. By the just completed Step 2, the function y_- has all the properties (i)–(iv) listed in Proposition 3.1 for x_-. Letting now $y_+ := x_{+,\nu}$, observe that the function y_+ too has the same four properties (the proof of these properties of $y_+ = x_{+,\nu}$ is practically the same as the corresponding part of the proof of Proposition 3.1).

Observe next that (cf. Proposition 3.25)

$$\int_0^{m_\nu} \frac{\mathrm{d}h}{y_+(h)} = \nu\big((0,\infty)\big) = \int_{(0,\infty)\times[0,1]} \nu(\mathrm{d}x)\,\mathrm{d}u.$$

Similarly to (4.19) but using (4.21) instead of (4.18), one also has

$$\int_0^{m_\nu} \frac{\mathrm{d}h}{y_-(h)} = \int_{(0,\infty)\times[0,1]} \frac{x}{\mathsf{s}(x,u)}\,\nu(\mathrm{d}x)\,\mathrm{d}u.$$

Hence, by condition (k), $\int_0^{m_\nu} \frac{\mathrm{d}h}{y_+(h)} + \int_0^{m_\nu} \frac{\mathrm{d}h}{-y_-(h)} \leqslant 1$.

It follows now by Proposition 3.26 that there exists a unique zero-mean probability measure μ such that $x_{+,\mu} = y_+$ and $x_{-,\mu} = y_-$. Since $x_{+,\mu} = y_+ = x_{+,\nu}$, in view of (3.1) one has

$$G_\mu = G_\nu \quad \text{on} \quad [0,\infty] \tag{4.26}$$

and hence, by condition (e), $\mu_+ = \nu$. Also, (4.26) implies that $\tilde{G}_\mu = \tilde{G}_\nu$ on $[0,\infty]\times[0,1]$. Now, by virtue of equalities (4.21), $y_- = x_{-,\mu}$, and (2.6), one has $\mathsf{r}_+ = \mathsf{s}$ (again for $\mathsf{r} := \mathsf{r}_\mu$).

Finaly, it remains to prove the uniqueness of μ given μ_+ and r_+. By Step 1 above, the function $x_{-,\mu}$ is uniquely determined by the condition

$$\mathsf{r}_+(x,u) = x_{-,\mu_+}\big(\tilde{G}_{\mu_+}(x,u)\big) \quad \text{for all} \quad (x,u) \in [0,\infty]\times[0,1]$$

(cf. (4.18) and (4.21)). Also, the function $x_{+,\mu}$ is uniquely determined by $\tilde{G}_{\mu_+}$. It remains to refer to the uniqueness part of the statement of Proposition 3.27. □

Proof of Proposition 3.28.

$\boxed{\text{Checking (I)} \Longrightarrow \text{(II)}}$ To prove this implication, assume that condition (I) holds. Then conditions 3.27(II)(a)–(d) follow by Proposition 3.27. Let us now prove that conditions 3.28(II)(i'), (j'), (k') also hold. Let $b := b_\mathsf{s}$ and $a := a_\mathsf{s}$.

$\boxed{\text{Checking (i'):}}$ By (already established) condition 3.27(II)(b) and definitions (3.47) and (3.48), $(b,\infty]\times[0,1] \subseteq M_\mathsf{s}(-\infty)$ while $M_\mathsf{s}(-\infty) \cap \big([0,b)\times[0,1]\big) = \emptyset$. Let us consider the cases $b=\infty$ and $b<\infty$ separately.

Case 1: $b = \infty$. Then $\mathsf{s} > -\infty$ on $[0,\infty)\times[0,1]$. Omitting the subscript $_\mu$ everywhere, recalling (2.6), and using the fact that $\mu(\{\infty\}) = 0$, one has $\mathsf{s}(\infty,u) = \mathsf{r}(\infty,u) = x_-\big(\tilde{G}(\infty,0)\big) = x_-\big(\tilde{G}(\infty,1)\big) = \mathsf{r}(\infty,1) = \mathsf{s}(\infty,1)$ for all $u \in [0,1]$, so that either $M_\mathsf{s}(-\infty) = \{\infty\}\times[0,1] = [b,\infty]\times[0,1]$ or $M_\mathsf{s}(-\infty) = \emptyset$, depending on whether $\mathsf{s}(\infty,1) = -\infty$ or not. Thus, (i') holds in Case 1.

Case 2: $b \in [0,\infty)$. Then, by definition (3.48), convention $b = b_\mathsf{s}$, and monotonicity condition 3.27(II)(b), there exists a nonincreasing sequence (x_n) in $[0,\infty)$ such that $(x_n,1) \in M_\mathsf{s}(-\infty)$ for all n and $x_n \searrow b$. So, for all n one has $\mathsf{s}(x_n,1) = -\infty$ and, by condition 3.27(II)(i), $\tilde{G}(x_n,1) = m$ or, equivalently, $G(x_n) = m$. Hence, by the right continuity of G on $[0,\infty)$, $G(b) = m$, and so, for all

n one has $G(b) = G(x_n)$, which (together with $\mathsf{s} = \mathsf{r}_+$) in turn yields $\mathsf{s}(b,1) = \mathsf{s}(x_n,1) = -\infty$. In particular, in view of condition 3.27(II)(a), $b \neq 0$, so that $b \in (0,\infty)$. Now let us consider separately the two subcases of Case 2, depending on whether $\mathsf{s}(b,u) > -\infty$ for all $u \in [0,1)$.

Subcase 2.1: $\mathsf{s}(b,u) > -\infty$ *for all* $u \in [0,1)$. Then $M_\mathsf{s}(-\infty) = \{(b,1)\} \cup (b,\infty] \times [0,1]$, so that (i') holds.

Subcase 2.2: $\mathsf{s}(b,u_b) = -\infty$ *for some* $u_b \in [0,1)$. Then, by 3.27(II)(i), $\tilde{G}(b,u_b) = m$. Now, if it were true that $\tilde{G}(b,0) < m$, then it would follow that $m = \tilde{G}(b,u_b) = (1-u_b)\tilde{G}(b,0) + u_b\tilde{G}(b,1) < m$, since $0 \leqslant u_b < 1$. This contradiction shows that $\tilde{G}(b,0) = m$ and hence for all $u \in [0,1]$ one has $m \geqslant \tilde{G}(b,u) \geqslant \tilde{G}(b,0) = m$, and so, $\tilde{G}(b,u) = m = \tilde{G}(b,u_b)$. This implies $\mathsf{s}(b,u) = \mathsf{s}(b,u_b) = -\infty$ for all $u \in [0,1]$. So, $M_\mathsf{s}(-\infty) = [b,\infty] \times [0,1]$, and again (i') holds.

This completes the checking of (i').

$\boxed{\text{Checking (j'):}}$ Here, note first that, by (a), $(0,u) \in M_\mathsf{s}(0)$ for all $u \in [0,1]$. Also, $[0,a) \times [0,1] \subseteq M_\mathsf{s}(0)$ while $M_\mathsf{s}(0) \cap \big((a,\infty] \times [0,1]\big) = \emptyset$. Let us consider the cases $a = 0$ and $a > 0$ separately.

Case 1: $a = 0$. Then $\mathsf{s} < 0$ on $(0,\infty] \times [0,1]$. So, $M_\mathsf{s}(0) = \{0\} \times [0,1] = [0,a] \times [0,1]$, and (j') holds.

Case 2: $a \in (0,\infty]$. Then by definition (3.48), convention $a = a_\mathsf{s}$, and monotonicity condition 3.27(II)(b), there exists a nondecreasing sequence (x_n) in $(0,\infty]$ such that $(x_n,0) \in M_\mathsf{s}(0)$ for all n and $x_n \nearrow a$. Therefore, for all n one has $\mathsf{s}(x_n,0) = 0$ and, by condition 3.27(II)(c), $\mathsf{s}(a,0) = 0$. Now let us consider separately the two subcases of Case 2, depending on whether $\mathsf{s}(a,u) < 0$ for all $u \in (0,1]$.

Subcase 2.1: $\mathsf{s}(a,u) < 0$ *for all* $u \in (0,1]$. Then $M_\mathsf{s}(0) = [0,a) \times [0,1] \cup \{(a,0)\}$. At that, $a < \infty$ – because $\mu(\{\infty\}) = 0$, and so, by condition 3.27(II)(g), $a = \infty$ would imply $\mathsf{s}(a,1) = \mathsf{s}(a,0) = 0$, which would contradict the definition of Subcase 2.1. Also, $a > 0$, since $\mathsf{s}(a,u) = 0$ for $a = 0$ and all $u \in [0,1]$, by condition 3.27(II)(a). Thus, (j') holds.

Subcase 2.2: $\mathsf{s}(a,u_a) = 0$ *for some* $u_a \in (0,1]$. Then, by 3.27(II)(j), $\tilde{G}(a,u_a) = 0$. Now, if it were true that $\tilde{G}(a,1) > 0$, then it would follow that $0 = \tilde{G}(a,u_a) = (1-u_a)\tilde{G}(a,0) + u_a\tilde{G}(a,1) > 0$, since $0 < u_a \leqslant 1$. This contradiction shows that $\tilde{G}(a,1) = 0$ and hence for all $u \in [0,1]$ one has $\tilde{G}(a,u) = 0$, and so, $\mathsf{s}(a,u) = 0$. This implies $M_\mathsf{s}(0) = [0,a] \times [0,1]$, and again (j') holds.

This completes the checking of (j').

$\boxed{\text{Checking (k'):}}$ Suppose that indeed $\mathsf{s}(a,1) \neq \mathsf{s}(a,0)$. Then, by 3.27(II)(a,g,b), one has $a\,\mu(\{a\}) > 0$ and $\mathsf{s}(a,1) < \mathsf{s}(a,0) \leqslant 0$. So,

$$\int_{[0,1]} \frac{a}{-\mathsf{s}(a,u)}\,\mu(\{a\})\,\mathrm{d}u \leqslant \int_{(0,\infty)\times[0,1]} \Big(1 - \frac{x}{\mathsf{s}(x,u)}\Big)\,\mu(\mathrm{d}x)\,\mathrm{d}u \leqslant 1$$

by 3.27(II)(k), whence (k') follows.

Thus, implication (I) $\Longrightarrow$ (II) is proved.

$\boxed{\text{Checking (II)} \Longrightarrow \text{(I)}}$ To prove this implication, assume that condition (II) indeed holds. Then, in view of (i'), b is never 0 $\big($in particular, $b = \infty$ if $M_\mathsf{s}(-\infty) = \emptyset\big)$. Also, by 3.27(II)(b), $a \leqslant b$. So, one has one of the following three cases: $0 \leqslant a < b \leqslant \infty$, $a = b = \infty$, and $0 < a = b < \infty$, which we shall consider separately.

Case 1: $0 \leqslant a < b \leqslant \infty$. Then consider the function φ defined on the set $R_a := [a,\infty) \cap (0,\infty)$ by the formula

$$\varphi(x) := \int_0^1 \Big(1 + x - \frac{x}{\mathsf{s}(x,u)}\Big)\,\mathrm{d}u; \tag{4.27}$$

here, $\frac{x}{\mathsf{s}(x,u)} := 0$ if $\mathsf{s}(x,u) = -\infty$; also, $\varphi(x) := \infty$ if $\mathsf{s}(x,u) = 0$ for all $u \in [0,1]$. One can see that this definition is correct; moreover, $\varphi(x) \in (1,\infty]$ for all $x \in R_a$, and $\varphi(x) = \infty$ for some $x \in R_a$ only if $x = a$ and $\mathsf{s}(a,1) = \mathsf{s}(a,0)$. Indeed, if $x > a$ then, by 3.27(II)(b,a) and (3.48), $\mathsf{s}(x,u) \leqslant \mathsf{s}(x,0) < 0$ for all $u \in [0,1]$, so that $0 < \int_0^1 \frac{x}{-\mathsf{s}(x,u)}\,\mathrm{d}u < \infty$. In particular, for all $x \in (a,b)$ one has $\varphi(x) \in (1,\infty)$. It also follows that, if $x \in R_a$ and $\mathsf{s}(x,u) = 0$ for some but not all $u \in [0,1]$, then necessarily $x = a$ and, by 3.27(II)(b), $\mathsf{s}(x,1) \neq \mathsf{s}(x,0)$, so that (k') implies that the integral $\int_0^1 \frac{x}{-\mathsf{s}(x,u)}\,\mathrm{d}u$ is finite $\big($and, by 3.27(II)(a,b), it is also strictly positive$\big)$. Let also $\varphi(x) := 0$ for $x \in \mathbb{R} \setminus R_a$.

Now one is prepared to introduce the measure ν_1 by the formula

$$\nu_1(A) := \int_{A\cap(a,b)} \frac{p_1(x)\,\mathrm{d}x}{\varphi(x)}$$

for all $A \in \mathscr{B}(\mathbb{R})$, where p_1 is any probability density function that is strictly positive on (a,b) and zero elsewhere. Recall that $\varphi(x) \in (1,\infty)$ for all $x \in (a,b)$. So, the measure ν_1 is finite, nonzero, nonnegative, and non-atomic (even more, it is absolutely continuous, with respect to the Lebesgue measure). Moreover, $\int_{\mathbb{R}} \varphi\,\mathrm{d}\nu_1 = \int_{(a,b)} p_1(x)\,\mathrm{d}x = 1$, $0 < \nu_1(\mathbb{R}) = \nu_1\big((0,\infty)\big) < 1$, and $\operatorname{supp}\nu_1 = \mathbb{R} \cap [a,b]$.

Next, introduce the (possibly empty) set

$$D := \{x \in (0,\infty] \colon \mathsf{s}(x,1) \neq \mathsf{s}(x,0)\}. \tag{4.28}$$

Observe that

$$D \subseteq [a,b]. \tag{4.29}$$

Indeed, if $x > b$ then, by 3.28(i'), $(x,u) \in M_{\mathsf{s}}(-\infty)$ for all $u \in [0,1]$, whence $\mathsf{s}(x,0) = -\infty = \mathsf{s}(x,1)$, and so, $x \notin D$. Similarly, if $0 \leqslant x < a$ then, by 3.28(j'), $(x,u) \in M_{\mathsf{s}}(0)$ for all $u \in [0,1]$, whence $\mathsf{s}(x,0) = 0 = \mathsf{s}(x,1)$, and so, $x \notin D$.

Observe also that the set D is at most countable; this follows because, for any x and y in D such that $x < y$, one has, by 3.27(b), $\mathsf{s}(y,1) < \mathsf{s}(y,0) \leqslant \mathsf{s}(x,1) < \mathsf{s}(x,0)$, so that the open intervals $\big(\mathsf{s}(x,1),\mathsf{s}(x,0)\big)$ and $\big(\mathsf{s}(y,1),\mathsf{s}(y,0)\big)$ are disjoint. So, there exists a function $p_2 \colon D \to (0,\infty)$ such that $\sum_{x\in D} p_2(x) = 1$ (unless $D = \emptyset$). Take now any such function p_2 and introduce the finite nonnegative discrete measure ν_2 by the formula

$$\nu_2(A) := \sum_{x\in A\cap D} \frac{p_2(x)}{\varphi(x)} \tag{4.30}$$

for all $A \in \mathscr{B}(\mathbb{R})$, where φ is still given by (4.27). Since $D \subseteq R_a$ and $\varphi(x) = \infty$ for some $x \in R_a$ only if $x = a$ and $\mathsf{s}(a,1) = \mathsf{s}(a,0)$, it follows that $1 < \varphi(x) < \infty$ for all $x \in D$. Therefore, definition (4.30) is correct; moreover, $\nu_2(\{x\}) > 0$ for all $x \in D$, while $0 \leqslant \nu_2(\mathbb{R}) = \nu_2\big((0,\infty)\big) < 1$. Furthermore, $\int_{\mathbb{R}} \varphi\,\mathrm{d}\nu_2 = \sum_{x\in D} p_2(x) = 1$ (unless $D = \emptyset$, in which case ν_2 is the zero measure).

Let $\psi(x) := \int_0^1 \left(1 - \frac{x}{\mathsf{s}(x,u)}\right) \mathrm{d}u$ for $x \in R_a$, so that $1 \leqslant \psi \leqslant \varphi$ on R_a, and hence $0 < \nu_1(\mathbb{R}) = \nu_1(R_a) \leqslant \int_{R_a} \psi \, \mathrm{d}(\nu_1 + \nu_2) \leqslant \int_{R_a} \varphi \, \mathrm{d}(\nu_1 + \nu_2) \leqslant 2$. Therefore, there exists a finite strictly positive constant c such that for the measure

$$\nu := c(\nu_1 + \nu_2) \tag{4.31}$$

one has $\int_{R_a} \psi \, \mathrm{d}\nu = 1$, so that condition 3.27(II)(k) is satisfied. Clearly, conditions 3.27(II)(e,f) will then be satisfied as well.

To complete the consideration of Case 1 in the proof of implication (II) $\Longrightarrow$ (I), it suffices to show that conditions 3.27(II)(g), (h), (i), (j) hold, for any choice of $c \in (0, \infty)$ in (4.31).

Checking 3.27(II)(g) Suppose that $x \in (0, \infty]$ and $\mathsf{s}(x,1) \neq \mathsf{s}(x,0)$, so that $x \in D$ and hence, by (4.31), $\nu(\{x\}) = c\,\nu_2(\{x\}) > 0$. This verifies 3.27(II)(g).

Checking 3.27(II)(h) Suppose that $0 \leqslant x < y \leqslant \infty$ and $\nu\big((x,y)\big) = 0$. Then, by (4.31), $\nu_1\big((x,y)\big) = 0$. Since $\operatorname{supp} \nu_1 = \mathbb{R} \cap [a,b]$, it follows that either $0 \leqslant x < y \leqslant a$ or $b \leqslant x < y \leqslant \infty$. By 3.28(II)(j'), $M_\mathsf{s}(0) \ni (a,0)$, whence $\mathsf{s}(a,0) = 0$, and so, $0 \leqslant x < y \leqslant a$ will imply $\mathsf{s}(y,0) = 0$ and $\mathsf{s}(x,1) = 0$, and hence $\mathsf{s}(x,1) = \mathsf{s}(y,0)$. In the other case, when $b \leqslant x < y \leqslant \infty$, it follows that $b < \infty$, and so, by 3.28(II)(i'), $(b,1) \in M_\mathsf{s}(-\infty)$; therefore, $\mathsf{s}(b,1) = -\infty$, and hence $\mathsf{s}(x,1) = -\infty = \mathsf{s}(y,0)$. This completes the verification of 3.27(II)(h).

Checking 3.27(II)(i) Take any $(x,u) \in [0,\infty) \times [0,1]$ such that $\mathsf{s}(x,u) = -\infty$. Then, by 3.28(II)(i'), $x \in [b, \infty)$. If, moreover, $x > b$ then, by (4.29), $[x,\infty) \cap D = \emptyset$, whence $\nu\big([x,\infty)\big) = c\,\nu_1\big([x,\infty)\big) + c\,\nu_2\big([x,\infty)\big) = 0$; so, $m_\nu \geqslant \tilde{G}_\nu(x,u) \geqslant G_\nu(x-) = m_\nu - \int_{[x,\infty)} z\,\nu(\mathrm{d}z) = m_\nu$, and we conclude that indeed $\tilde{G}_\nu(x,u) = m_\nu$.

Let now $x = b$, so that the assumption $\mathsf{s}(x,u) = -\infty$ becomes $\mathsf{s}(b,u) = -\infty$; this admits only two possible forms of $M_\mathsf{s}(-\infty)$ of the three ones listed in condition 3.28(II)(i'). Accordingly, let us consider the following two subcases.

Subcase i1: $M_\mathsf{s}(-\infty) = \{(b,1)\} \cup \big((b,\infty] \times [0,1]\big)$. Then necessarily $u = 1$, and so, $\tilde{G}_\nu(x,u) = \tilde{G}_\nu(b,1) = G_\nu(b) = m_\nu - \int_{(b,\infty)} z\,\nu(\mathrm{d}z) = m_\nu$, since $\operatorname{supp}\nu \subseteq [a,b]$.

Subcase i2: $M_\mathsf{s}(-\infty) = [b,\infty] \times [0,1]$. Then $\mathsf{s}(b,0) = -\infty = \mathsf{s}(b,1)$. So, $b \notin D$ and hence $\nu_2(\{b\}) = 0$, which yields $\nu(\{b\}) = 0$. Therefore, $m_\nu \geqslant \tilde{G}_\nu(x,u) = \tilde{G}_\nu(b,u) \geqslant G_\nu(b-) = m_\nu - \int_{[b,\infty)} z\,\nu(\mathrm{d}z) = m_\nu$, since $\operatorname{supp}\nu \subseteq [a,b]$ and $\nu(\{b\}) = 0$.

Thus, in both subcases one has $\tilde{G}_\nu(x,u) = m_\nu$. This completes the verification of 3.28(II)(i).

Checking 3.27(II)(j) This is similar to checking (i), so we shall provide fewer details. Take any $(x,u) \in (0,\infty] \times [0,1]$ such that $\mathsf{s}(x,u) = 0$. Then, by 3.28(II)(j'), $x \in [0,a]$. If, moreover, $x < a$ then $\nu\big([0,x]\big) = 0$; so, $0 \leqslant \tilde{G}_\nu(x,u) \leqslant G_\nu(x) = \int_{(0,x]} z\,\nu(\mathrm{d}z) = 0$, and we conclude that indeed $\tilde{G}_\nu(x,u) = 0$.

Let now $x = a$, so that $\mathsf{s}(a,u) = 0$. Consider the two possible subcases.

Subcase j1: $M_\mathsf{s}(0) = \big([0,a) \times [0,1]\big) \cup \{(a,0)\}$. Then $u = 0$, and so, $\tilde{G}_\nu(x,u) = \tilde{G}_\nu(a,0) = G_\nu(a-) = \int_{(0,a)} z\,\nu(\mathrm{d}z) = 0$.

Subcase j2: $M_\mathsf{s}(0) = [0,a] \times [0,1]$. Then $\mathsf{s}(a,0) = 0 = \mathsf{s}(a,1)$. So, $a \notin D$ and hence $\nu_2(\{a\}) = 0$ and $\nu(\{a\}) = 0$. Therefore, $0 \leqslant \tilde{G}_\nu(x,u) = \tilde{G}_\nu(a,u) \leqslant G_\nu(a) = \int_{(0,a]} z\,\nu(\mathrm{d}z) = 0$, since $\operatorname{supp}\nu \subseteq [a,b]$ and $\nu(\{a\}) = 0$.

Thus, in both subcases one has $\tilde{G}_\nu(x,u) = 0$. This completes the verification of (j), and thus the entire proof of implication (II) $\Longrightarrow$ (I) in Case 1.

Case 2: $a = b = \infty$. Then, by 3.28(II)(j'), $M_{\mathsf{s}}(0) = [0,\infty]\times[0,1]$ and $M_{\mathsf{s}}(-\infty) = \emptyset$. So, the function s is the constant 0. Then, letting $\nu = \delta_0$, one sees that conditions 3.27(II)(e)–(h), (k) are all trivial, with $m_\nu = 0$ and $\tilde{G}_\nu \equiv 0$, whence conditions 3.27(II)(i,j) are also trivial.

Case 3: $0 < a = b < \infty$. Then 3.28(II)(i') and 3.28(II)(j') admit only the following possibility: $M_{\mathsf{s}}(0) = \big([0,a)\times[0,1]\big)\cup\{(a,0)\}$ and $M_{\mathsf{s}}(-\infty) = \{(a,1)\}\cup\big((a,\infty]\times[0,1]\big)$. Let now $\nu := c\,\delta_a$, where $c \in (0,\infty)$. Then conditions 3.27(II)(e,f) are trivial, with $m_\nu = ca$. By 3.28(II)(k'), one can satisfy (k) by taking a small enough c.

It remains to check conditions 3.27(II)(g), (h), (i), (j) – in Case 3.

$\boxed{\text{Checking 3.27(II)(g)}}$ Assume here that $\nu(\{x\}) = 0$ and $x \in (0,\infty]$. Then either $x \in (0,a)$ or $x \in (a,\infty]$. If $x \in (0,a)$ then $\mathsf{s}(x,1) = 0 = \mathsf{s}(x,0)$. If $x \in (a,\infty]$ then $\mathsf{s}(x,1) = -\infty = \mathsf{s}(x,0)$. This verifies 3.27(II)(g).

$\boxed{\text{Checking 3.27(II)(h)}}$ Assume here that $\nu\big((x,y)\big) = 0$ and $0 \leqslant x < y \leqslant \infty$. Then either $0 \leqslant x < y \leqslant a$ or $a \leqslant x < y \leqslant \infty$. If $0 \leqslant x < y \leqslant a$ then $\mathsf{s}(x,1) = 0 = \mathsf{s}(y,0)$, since both points $(x,1)$ and $(y,0)$ are in $M_{\mathsf{s}}(0) = \big([0,a)\times[0,1]\big)\cup\{(a,0)\}$. Similarly, if $a \leqslant x < y \leqslant \infty$ then $\mathsf{s}(x,1) = -\infty = \mathsf{s}(y,0)$, since both points $(x,1)$ and $(y,0)$ are in $M_{\mathsf{s}}(-\infty) = \{(a,1)\}\cup\big((a,\infty]\times[0,1]\big)$. This verifies 3.27(II)(h).

The verification of conditions 3.27(II)(i), (j) in Case 3 is similar to that in Case 1, but simpler.

Thus, the proof of implication (II) $\Longrightarrow$ (I) and thereby the entire proof of Proposition 3.28 is complete. ☐

Proof of Proposition 3.29.

$\boxed{\text{Checking (I)} \Longrightarrow \text{(II).}}$ Assume that condition (I) holds. By Proposition 3.28, here it suffices to check that condition 3.29(II)(u) holds. To obtain a contradiction, assume that 3.29(II)(u) is false. Then, by property 3.27(II)(b), there exist some x and y such that

$$0 \leqslant x < y \leqslant \infty \text{ but } -\infty < \mathsf{s}(x,1) = \mathsf{s}(y,0) < 0.$$

Hence, by (3.48), $a \leqslant x < y \leqslant b$; in particular, $a < b$. Take now any $\tilde{x}$ and $\tilde{y}$ such that $x < \tilde{x} < \tilde{y} < y$. Then, again by property 3.27(II)(b),

$$\mathsf{s}(x,1) \geqslant \mathsf{s}(\tilde{x},1) \geqslant \mathsf{s}(\tilde{y},0) \geqslant \mathsf{s}(y,0) = \mathsf{s}(x,1), \tag{4.32}$$

whence $\mathsf{s}(\tilde{x},1) = \mathsf{s}(\tilde{y},0)$. Also, $\mathsf{s}(x,1) \geqslant \mathsf{s}(\tilde{x},0) \geqslant \mathsf{s}(\tilde{x},1) \geqslant \mathsf{s}(y,0) = \mathsf{s}(x,1)$, and so, $\mathsf{s}(\tilde{x},1) = \mathsf{s}(\tilde{x},0)$ for all $\tilde{x} \in (x,y)$; that is, $(x,y)\cap D = \emptyset$, where D is the set defined by (4.28). So, there exist $\tilde{x}$ and $\tilde{y}$ such that

$$0 \leqslant a < \tilde{x} < \tilde{y} < b \leqslant \infty, \quad -\infty < \mathsf{s}(\tilde{x},1) = \mathsf{s}(\tilde{y},0) < 0, \quad [\tilde{x},\tilde{y}]\cap D = \emptyset. \tag{4.33}$$

Now, fix any such $\tilde{x}$ and $\tilde{y}$ and – along with the measures ν_1, ν_2, and ν constructed in the proof of Proposition 3.28 – consider the measures $\tilde{\nu}_1$ and $\tilde{\nu}$ defined by the formulas

$$\tilde{\nu}_1(A) := \nu_1\big(A\setminus(\tilde{x},\tilde{y})\big) \quad \text{for all } A \in \mathscr{B}(\mathbb{R}) \quad \text{and} \quad \tilde{\nu} := c\,(\tilde{\nu}_1 + \nu_2),$$

where c again is a finite positive constant.

Then one can verify that properties 3.27(II)(e)–(k) hold with $\tilde{\nu}$ in place of ν. This verification is quite similar to that done when checking implication (II) $\Longrightarrow$ (I) of Proposition 3.28, with the only (if any) substantial difference being in the verification of (h). There, another case when the conjuncture of conditions $0 \leqslant x < y \leqslant \infty$ and $\nu\big((x,y)\big) = 0$ occurs is $\tilde{x} \leqslant x < y \leqslant \tilde{y}$. But then too (cf. (4.32)), it easily follows that $\mathsf{s}(x,1) = \mathsf{s}(y,0)$.

So, by Proposition 3.27, there exist zero-mean probability measures μ and $\tilde{\mu}$ such that for $\mathsf{r} := \mathsf{r}_\mu$ and $\tilde{\mathsf{r}} := \mathsf{r}_{\tilde{\mu}}$ one has $\mu_+ = \nu$, $\tilde{\mu}_+ = \tilde{\nu}$, and $\mathsf{r}_+ = \tilde{\mathsf{r}}_+ = \mathsf{s}$. Moreover, by the last condition in (4.33),

$$\tilde{\nu}\big((\tilde{x},\tilde{y})\big) = 0. \tag{4.34}$$

Fix now any $x \in (\tilde{x},\tilde{y})$ and let, for brevity, $\hat{x} := \hat{x}_\mu(x,1)$. If $\hat{x} < x$ then $\mu\big((\hat{x},x)\big) = \nu\big((\hat{x},x)\big) \geqslant c\,\nu_1\big((\hat{x},x)\big) > 0$, since $c > 0$ and $x \in (a,b) = \text{interior}(\operatorname{supp}\nu_1)$; this contradicts properties (i) and (ii)(c) listed in Proposition 3.8. So, by Proposition 3.11, $x = \hat{x}_\mu(x,1) = \mathsf{r}\big(\mathsf{r}(x,1),v\big)$ for some $v \in [0,1]$. But, by condition (I) of Proposition 3.29, $\mathsf{r} = \tilde{\mathsf{r}}$, and so, $x = \tilde{\mathsf{r}}\big(\tilde{\mathsf{r}}(x,1),v\big)$. Therefore, by (2.6),

$$x = x_{+,\tilde{\mu}}(h) = x_{+,\tilde{\nu}}(h) \quad \text{for some } h \in [0,m_{\tilde{\nu}}]. \tag{4.35}$$

Next, introduce $h_0 := G_{\tilde{\nu}}(x)$. Then $G_{\tilde{\nu}}(\tilde{x}) = h_0$ $\big($because, by (4.34), $\tilde{\nu}\big((\tilde{x},x]\big) = 0\big)$. So, for each $h \in (h_0,m_{\tilde{\nu}}]$, one has $G_{\tilde{\nu}}(x) < h$, whence, by (3.1), $x < x_{+,\tilde{\nu}}(h)$. On the other hand, again by (3.1) and the mentioned relation $G_{\tilde{\nu}}(\tilde{x}) = h_0$, for each $h \in [0,h_0]$ one has $G_{\tilde{\nu}}(\tilde{x}) \geqslant h$, whence, again by (3.1), $\tilde{x} \geqslant x_{+,\tilde{\nu}}(h)$, and thus $x > x_{+,\tilde{\nu}}(h)$. So, $x < x_{+,\tilde{\nu}}(h)$ for each $h \in (h_0,m_{\tilde{\nu}}]$, and $x > x_{+,\tilde{\nu}}(h)$ for each $h \in [0,h_0]$. This is a contradiction with (4.35). Thus, condition 3.29(II)(u) follows, so that implication (I) $\Longrightarrow$ (II) of Proposition 3.29 is verified.

$\boxed{\text{Checking (II)} \Longrightarrow \text{(I)}}$ Assume that condition (II) holds. We have to show that there exists a *unique* function r such that $\mathsf{r}_+ = \mathsf{s}$ and r coincides with the reciprocating function r_μ of some zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$. The existence here follows immediately from Proposition 3.28. To verify implication (II) $\Longrightarrow$ (I), it remains to prove the uniqueness of r. We shall do it in steps.

$\boxed{\textit{Step 1.}}$ Here we shall prove that the values of $\mathsf{r}(z,v)$ are uniquely determined for all $(z,v) \in (-\infty,0) \times (0,1]$ such that z is in the image, say S, of the set $(0,\infty] \times [0,1]$ under the mapping s. More specifically, we shall prove at this step that, if $z = \mathsf{s}(x,u) \in (-\infty,0)$ for some $(x,u) \in (0,\infty] \times [0,1]$, then $\mathsf{r}(z,v) = x$ for all $v \in (0,1]$. Toward that end, fix any any $v \in (0,1]$ and any $(x,u) \in (0,\infty] \times [0,1]$ such that

$$z := \mathsf{s}(x,u) \in (-\infty,0); \tag{4.36}$$

then one also has $\mathsf{r}(x,u) = \mathsf{s}(x,u) = z$. Next, introduce

$$y := \mathsf{r}(z,v).$$

Then $y \in [0,\infty]$ and, by Proposition 3.11, there exists some $w \in [0,1]$ such that $\mathsf{r}(y,w) = \mathsf{r}\big(\mathsf{r}(z,v),w\big) = \hat{x}(z,v)$. On the other hand, $z = \mathsf{r}(x,u) = x_-(h)$ for $h := \tilde{G}(x,u)$. So, by property (iv)(f) of Proposition 3.8 and because $v \neq 0$, one has $\hat{x}(z,v) = z$. So, recalling that $\mathsf{r}(y,w) = \hat{x}(z,v)$, one has

$$z = \mathsf{r}(y,w) = \mathsf{s}(y,w). \tag{4.37}$$

Next, consider two cases: $y > x$ and $x > y$, to show that either one effects a contradiction.

Case 1: $y > x$, so that $0 \leqslant x < y \leqslant \infty$. Then, by condition 3.29(II)(u), one of the following two subcases must take place: $\mathsf{s}(x,1) > \mathsf{s}(y,0)$ or $\mathsf{s}(x,1) = \mathsf{s}(y,0) \in \{-\infty, 0\}$.

Subcase 1.1: $\mathsf{s}(x,1) > \mathsf{s}(y,0)$. Then, by property 3.27(II)(b) of s, (4.36), and (4.37), one has

$$\mathsf{s}(x,1) \leqslant \mathsf{s}(x,u) = z = \mathsf{s}(y,w) \leqslant \mathsf{s}(y,0) < \mathsf{s}(x,1), \tag{4.38}$$

which is a contradiction.

Subcase 1.2: $\mathsf{s}(x,1) = \mathsf{s}(y,0) \in \{-\infty, 0\}$. Then the "non-strict" version of (4.38), with the sign $<$ replaced by $\leqslant$, still holds, whence $z \in \{-\infty, 0\}$, which contradicts the assumption $(z,v) \in (-\infty,0) \times (0,1]$ made above for Step 1.

Case 2: $x > y$. This case is treated similarly to Case 1, with the roles of the pairs (x,u) and (y,w) interchanged.

From this consideration of Cases 1 and 2, it follows that $y = x$, that is, $\mathsf{r}(z,v) = x$. This completes Step 1.

$\boxed{\textit{Step 2.}}$ Here we shall prove that the values of $\mathsf{r}(z,v)$ are uniquely determined for all $(z,v) \in (-\infty,0) \times (0,1]$, whether or not z is in the image, S, of the set $(0,\infty] \times [0,1]$ under the mapping s. Toward that end, fix any $(z,v) \in (-\infty,0) \times (0,1]$ and introduce

$$L_z := \{x \in [0,\infty] : \mathsf{s}(x,0) \geqslant z\} \quad \text{and} \quad x_z := \sup L_z.$$

Note that $0 \in L_z$, so that $L_z \neq \emptyset$ and $x_z \in [0,\infty]$. By the monotonicity and left-continuity properties 3.27(II)(b,c), $L_z = [0, x_z]$, so that

$$\mathsf{s}(x,0) \geqslant z \ \forall x \in [0,x_z] \quad \text{and} \quad \mathsf{s}(x,0) < z \ \forall x \in (x_z,\infty].$$

Let next

$$u_z := \sup\{u \in [0,1] : \mathsf{s}(x_z,u) \geqslant z\}.$$

Then $u_z \in [0,1]$ (because $\mathsf{s}(x_z,0) \geqslant z$) and, by the monotonicity and left-continuity properties 3.27(II)(b,d),

$$\mathsf{s}(x_z,u) \geqslant z \ \forall u \in [0,u_z] \quad \text{and} \quad \mathsf{s}(x_z,u) < z \ \forall u \in (u_z,1].$$

Thus, in terms of the lexicographic order,

$$\mathsf{s}(x,u) \begin{cases} \geqslant z & \text{if } (0,0) \preccurlyeq (x,u) \preccurlyeq (x_z,u_z), \\ < z & \text{if } (x,u) \succ (x_z,u_z), \end{cases} \tag{4.39}$$

where, as usual, $(x,u) \preccurlyeq (y,v)$ means that either $(x,u) \prec (y,v)$ or $(x,u) = (y,v)$, and $(x,u) \succ (y,v)$ means that $(y,v) \prec (x,u)$.

In particular, $\mathsf{s}(x_z,u_z) \geqslant z$. So, by Step 1, w.l.o.g. we may, and shall, assume that

$$\check{z} := \mathsf{s}(x_z,u_z) > z. \tag{4.40}$$

Note that the pair (x_z,u_z) is uniquely determined by z and the function s, and hence so is $\check{z}$. Note also that (4.40) implies that $-\infty < \check{z} \leqslant 0$.

Therefore, to complete Step 2, it suffices to verify that $\mathsf{r}(z,v) = \mathsf{r}(\check{z},1)$ for all $v \in [0,1]$. Indeed, then one will have $\mathsf{r}(z,v) = 0$ if $\check{z} = 0$ and (by Step 1 and inequalities $-\infty < \check{z} \leqslant 0$) $\mathsf{r}(z,v) = x_z$ if $\check{z} \neq 0$.

So, to obtain a contradiction, assume that $\mathsf{r}(z,v) \neq \mathsf{r}(\check{z},1)$ for some $v \in [0,1]$. Then (cf. the monotonicity property 3.27(II)(b)), by (4.40), $\mathsf{r}(z,v) > \mathsf{r}(\check{z},1)$, and so,

$$\mathsf{r}(z,1) > \mathsf{r}(\check{z},1). \tag{4.41}$$

Let us now consider separately the only two possible cases: $\check{z} = 0$ and $-\infty < \check{z} < 0$.

Case 1: $\check{z} = 0$. Then $0 = \check{z} = \mathsf{s}(x_z, u_z)$, so that, by (4.39), 3.27(II)(b), and (3.47), $M_\mathsf{s}(0) = \{(x,u) \colon (0,0) \preccurlyeq (x,u) \preccurlyeq (x_z,u_z)\} = [0,x_z) \times [0,1] \cup \{x_z\} \times [0,u_z]$. Comparing this with property 3.28(II)(j'), one concludes that either $u_z = 0$ or $u_z = 1$. Let us show that either of these subcases effects a contradiction.

Subcase 1.1: $u_z = 0$. Then, in view of (4.39) and equalities $\mathsf{s}(x_z,u_z) = \check{z} = 0$, one has

$$\mathsf{s}\begin{cases} = 0 & \text{on } [0,x_z) \times [0,1] \cup \{(x_z,0)\}, \\ < z & \text{on } \{x_z\} \times (0,1] \cup (x_z,\infty] \times [0,1]. \end{cases} \tag{4.42}$$

In particular, $0 = \mathsf{s}(x_z,0) = x_-\big(G(x_z-)\big)$, where the functions x_- and G pertain to any given zero-mean probability measure μ such that $(\mathsf{r}_\mu)_+ = \mathsf{s}$. So, by the strict positivity property (iii) listed in Proposition 3.1, $G(x_z-) = 0$. Take now any $u \in (0,1]$; then, by (4.42), $z > \mathsf{s}(x_z,u) = x_-\big(\tilde{G}(x_z,u)\big)$ and hence, by (3.2), $G(z) < \tilde{G}(x_z,u)$. So, $0 \leqslant G(z) \leqslant \lim_{u\downarrow 0} \tilde{G}(x_z,u) = G(x_z-) = 0$, whence $G(z) = 0$ and $\mathsf{r}(z,1) = x_+\big(G(z)\big) = x_+(0) = 0$, which contradicts inequality (4.41), since $\mathsf{r}(\check{z},1) = \mathsf{r}(0,1) = 0$.

Subcase 1.2: $u_z = 1$. This subcase is similar to Subcase 1.1. Indeed, here

$$\mathsf{s}\begin{cases} = 0 & \text{on } [0,x_z] \times [0,1], \\ < z & \text{on } (x_z,\infty] \times [0,1]. \end{cases} \tag{4.43}$$

In particular, $0 = \mathsf{s}(x_z,1) = x_-\big(G(x_z)\big)$, whence $G(x_z) = 0$. Also, for all $(x,u) \in (x_z,\infty] \times [0,1]$, (4.43) (or even (4.39)) yields $z > \mathsf{s}(x,u) = x_-\big(\tilde{G}(x,u)\big)$ and hence, by (3.2), $G(z) < \tilde{G}(x,u) \leqslant G(x)$. So, $0 \leqslant G(z) \leqslant \lim_{x\downarrow x_z} G(x) = G(x_z) = 0$, whence $G(z) = 0$ and $\mathsf{r}(z,1) = 0$, which again contradicts inequality (4.41).

Case 2: $-\infty < \check{z} < 0$. Then, by Step 1, $\mathsf{r}(\check{z},1) = x_z$, so that, in view of (4.41), $\mathsf{r}(z,1) > \mathsf{r}(\check{z},1) = x_z$. So, by (2.6), $x_+\big(G(z)\big) > x_z$. Hence, by (3.1), $G(x_z) < G(z)$. On the other hand, just in the consideration of Subcase 1.2, one can see that $G(z) \leqslant G(x_z)$, which contradicts the just established inequality $G(x_z) < G(z)$.

All these contradictions demonstrate that indeed $\mathsf{r}(z,v) = \mathsf{r}(\check{z},1)$ for all $v \in [0,1]$, which completes Step 2 in the proof of implication (II) $\Longrightarrow$ (I) of Proposition 3.29.

Step 3. Here we shall conclude the proof of implication (II) $\Longrightarrow$ (I). By Step 2, the values of $\mathsf{r}(z,v)$ are uniquely determined for all $(z,v) \in (-\infty,0) \times (0,1]$. Let now μ_1 and μ_2 be any two zero-mean probability measures such that $(\mathsf{r}_{\mu_1})_+ = \mathsf{s} = (\mathsf{r}_{\mu_2})_+$, and let A_1 and A_2 be, respectively, the sets of all the atoms of μ_1 and μ_2. Then the set $A := A_1 \cup A_2$ is countable. Also, for all $z \in (-\infty,0) \setminus A$ one has $\mathsf{r}_{\mu_1}(z,0) = \mathsf{r}_{\mu_1}(z,1) = r_{\mu_2}(z,1) = r_{\mu_2}(z,0)$. On the other hand (cf. condition 3.27(II)(c)), for any reciprocating function r, the function $z \mapsto \mathsf{r}(z,0)$ is right-continuous on $[-\infty,0)$. Since the set $(-\infty,0) \setminus A$ is dense in a right neighborhood of any given point in $[-\infty,0)$, it follows that $\mathsf{r}_{\mu_1}(z,0) = r_{\mu_2}(z,0)$ for all $z \in [-\infty,0)$, which in turn also implies that for all $u \in [0,1]$ one has $\mathsf{r}_{\mu_1}(-\infty,u)=r_{\mu_2}(-\infty,u)$, since $\tilde G_\mu(-\infty,u)=G_\mu\big((-\infty)+\big)=\tilde G_\mu(-\infty,0)$ for any probability measure μ on $\mathscr{B}(\mathbb{R})$.

Thus, $\mathsf{r}_{\mu_1}(z,u)=r_{\mu_2}(z,u)$ for all $u\in[0,1]$ and all $z\in[-\infty,0)$. The same trivially holds for all $z\in[0,\infty)$ $\big($since $(\mathsf{r}_{\mu_1})_+=(\mathsf{r}_{\mu_2})_+=\mathsf{s}\big)$.

This completes the proof of implication (II) $\Longrightarrow$ (I) of Proposition 3.29.

It remains to prove the statement about $\operatorname{supp}(\mu_+)$. Assume that indeed either one of the two mutually equivalent conditions, (I) or (II), holds. Let μ be any zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$ such that $(\mathsf{r}_\mu)_+=\mathsf{s}$. For brevity, let us again omit subscript $_\mu$ everywhere and write a and b for a_s and b_s. Then for all $(x,u)\in(b,\infty)\times[0,1]$ one has $\mathsf{s}(x,u)=-\infty$, whence, by 3.27(II)(i), $G(x)=m$. So, $G(b)=G(b+)=m$ if $b<\infty$. On the other hand, if $b=\infty$, then $G(b)=G(\infty)=m$. Therefore, in all cases $G(b)=m=G(\infty)$, which yields $\mu\big((b,\infty)\big)=0$.

Similarly, for all $(x,u)\in[0,a)\times[0,1]$ one has $\mathsf{s}(x,u)=0$, whence, by 3.27(II)(j), $G(x)=0$. So, $G(a-)=0=G(0)$ if $a>0$. On the other hand, if $a=0$, then $G(a-)=G(0)=0$. Therefore, in all cases $G(a-)=0=G(0)$, which yields $\mu\big((0,a)\big)=0$.

Thus, $\operatorname{supp}(\mu_+)\subseteq[a,b]$. To obtain a contradiction, assume that $\operatorname{supp}(\mu_+)\neq[a,b]$. Then there exist x and y such that $a<x<y<b$ and $\mu\big((x,y)\big)=0$, whence $\nu\big((x,y)\big)=0$. So, by 3.27(II)(h), $\mathsf{s}(x,1)=\mathsf{s}(y,0)$. Consequently, by 3.29(II)(u), $\mathsf{s}(x,1)$ is either $-\infty$ or 0. But this contradicts condition (3.48), since $a<x<b$. This contradiction shows that indeed $\operatorname{supp}(\mu_+)=\mathbb{R}\cap[a_\mathsf{s},b_\mathsf{s}]$. The proof of Proposition 3.29 is now complete. □

Proof of Proposition 3.30. Note that, in the proof of implication (I) $\Longrightarrow$ (II) in Proposition 3.29, condition 3.30(II)(u) was used only in Step 1, where it was proved that, if $z=\mathsf{s}(x,u)\in(-\infty,0)$ for some $(x,u)\in[0,\infty]\times[0,1]$, then $\mathsf{r}(z,v)=x$ for all $v\in(0,1]$.

So, here it suffices to verify that, if μ is any zero-mean probability measure on $\mathscr{B}(\mathbb{R})$ such that μ_- is non-atomic and $(\mathsf{r}_\mu)_+=\mathsf{s}$, then for all $(x,u)\in[0,\infty]\times[0,1]$

$$\mathsf{r}\big(\mathsf{s}(x,u)\big)=\inf\{y\in[0,x]\colon\mathsf{s}(y,u)=\mathsf{s}(x,u)\};\tag{4.44}$$

here $\mathsf{r}:=\mathsf{r}_\mu$, and we write $\mathsf{r}(z)$ instead of $\mathsf{r}(z,v)$ for any $(z,v)\in[-\infty,0]\times[0,1]$, which is correct since $\mu_-=\mu|_{\mathscr{B}((-\infty,0))}$ is non-atomic and $\mathsf{r}(0,v)=0$ for all $v\in[0,1]$. To prove (4.44), take any $(x,u)\in[0,\infty]\times[0,1]$ and introduce

$$z:=\mathsf{s}(x,u),\quad h:=\tilde G(x,u),\quad E:=\{y\in[0,x]\colon\mathsf{s}(y,u)=z\};\tag{4.45}$$

again, the subscript $_\mu$ is omitted everywhere here. Then $z\in[-\infty,0]$ and $x\in E$. Also, by (2.6) and $\mathsf{r}_+=\mathsf{s}$,

$$z=x_-(h),\tag{4.46}$$

whence, by (3.6), $G(z+)\leqslant h\leqslant G(z)$, and so, $G(z)=h$ $\big($since $z\leqslant 0$ and μ_- is non-atomic, so that G is continuous on $[-\infty,0]\big)$. Therefore, $\mathsf{r}(z)=x_+\big(G(z)\big)=x_+(h)$. So, by (4.45), to prove (4.44) it suffices to check that $x_+(h)=\inf E$.

Take now any $y\in E$. Then, by (2.6), $x_-\big(\tilde G(y,u)\big)=\mathsf{r}(y,u)=\mathsf{s}(y,u)=z$. Hence, by (4.46) and (3.8), $\tilde G(y,u)=h$ $\big($because μ_- is non-atomic and hence $z\,\mu(\{z\})=0\big)$. So, by (2.5), $G(y)\geqslant h$ and, then by (3.1), $x_+(h)\leqslant y$ – for any $y\in E$. It follows that $x_+(h)\leqslant\inf E$. So, it remains to show that $x_+(h)\geqslant\inf E$. Assume the contrary: $x_+(h)<\inf E$. Then also $x_+(h)<x$, since $x\in E$. So, there exists some y_1 such that $x_+(h)<y_1<x$ and $y_1\notin E$, so that, by (4.45), $\mathsf{s}(y_1,u)\neq z=\mathsf{s}(x,u)$. But

$\mathsf{s}(y_1,u) \geqslant \mathsf{s}(x,u)$, because $0 < y_1 < x$ and the function $\mathsf{s} = \mathsf{r}_+$ is non-increasing, by (3.27)(II)(b). It follows that

$$\mathsf{s}(y_1,u) > z. \tag{4.47}$$

On the other hand, by (3.4), $G\big(x_+(h)\big) \geqslant h \geqslant G\big(x_+(h)-\big)$. So, there exists some $v \in [0,1]$ such that $h = \tilde{G}\big(x_+(h),v\big)$, whence $\mathsf{s}\big(x_+(h),v\big) = x_-(h) = z$. Also, again by (3.27)(II)(b), the condition $x_+(h) < y_1$ implies that $\mathsf{s}(y_1,u) \leqslant \mathsf{s}\big(x_+(h),v\big) = z$, which contradicts (4.47). □

Proof of Proposition 3.31.

$\boxed{\text{Checking (I)} \Longrightarrow \text{(II).}}$ Here it is assumed that condition (I) of Proposition 3.31 takes place. By Proposition 3.27, conditions 3.27(II)(a)–(c), (e), (f), (i), and (j) $\big($with $\mathsf{s}(x)$ and $G(x)$ in place of $\mathsf{s}(x,u)$ and $\tilde{G}(x,u)\big)$ will then hold. So, to complete the proof of implication (I) $\Longrightarrow$ (II), it remains to check conditions (h') and (k").

$\boxed{\text{Checking (h').}}$ By 3.27(II)(h), only implication $\Longleftarrow$ in place of $\Longleftrightarrow$ needs to be proved here. To obtain a contradiction, suppose that $0 \leqslant x < y \leqslant \infty$, $\nu\big((x,y)\big) > 0$, and $\mathsf{s}(x) = \mathsf{s}(y)$, so that $G(x) < G(y)$ and $x_-\big(G(x)\big) = x_-\big(G(y)\big) =: z \leqslant 0$ (for brevity, here we omit the subscript $_\mu$). Then $G(z) \geqslant G(y) > G(x) \geqslant G(z+0)$, by (3.1) and (3.6). So, $\mu(\{z\}) > 0$, which contradicts the condition that measure μ is non-atomic.

$\boxed{\text{Checking (k").}}$ The verification of this is the same as that of condition 3.27(II)(k), taking also into account in (4.20) that $\mu\big(\mathbb{R} \setminus \{0\}\big) = 1$, since μ is non-atomic.

$\boxed{\text{Checking (II)} \Longrightarrow \text{(I).}}$ Here it is assumed that condition (II) of Proposition 3.31 takes place. Then, by Proposition 3.27, there exists a unique zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$ whose reciprocating function $\mathsf{r} = \mathsf{r}_\mu$ satisfies the conditions $\mu_+ = \nu$ and $\mathsf{r}_+ = \mathsf{s}$. It remains to show that μ is non-atomic. To obtain a contradiction, suppose that $\mu(\{z\}) > 0$ for some $z \in \mathbb{R}$. Then $z \in (-\infty,0)$, since measure $\mu_+ = \nu$ is non-atomic. Introduce $h_1 := G(z+)$ and $h_2 := G(z)$. (Here as well, we omit the subscript $_\mu$.) Then $\mu(\{z\}) > 0$ implies that $0 \leqslant h_1 < h_2 \leqslant m$. Take any $h \in (h_1,h_2)$ and let $x := x_+(h)$ and $y := x_+(h_2)$. Then $0 \leqslant x \leqslant y \leqslant \infty$. Also, by (3.4), $G(x) = h$ and $G(y) = h_2$, since $\mu_+ = \nu$ is non-atomic. So, $G(x) < G(y)$, whence $\nu\big((x,y)\big) > 0$ and $x < y$. So, by (h'), $\mathsf{s}(y) < \mathsf{s}(x)$. On the other hand, $\mathsf{s}(x) = x_-\big(G(x)\big) = x_-(h)$ and $\mathsf{s}(y) = x_-\big(G(y)\big) = x_-(h_2) \geqslant z$, by (3.2). Also, taking any $w \in (z,0)$, one has $G(w) \leqslant G(z+) = h_1 < h$ and hence, again by (3.2), $x_-(h) < w$. This implies that $\mathsf{s}(x) = x_-(h) \leqslant z$. Thus, $z \leqslant \mathsf{s}(y) < \mathsf{s}(x) \leqslant z$, a contradiction. □

Proof of Proposition 3.32.

$\boxed{\text{Checking (I)} \Longrightarrow \text{(II).}}$ Here it is assumed that condition (I) of Proposition 3.32 takes place. By Proposition 3.31, conditions 3.27(II)(a)–(c) $\big($with $\mathsf{s}(x)$ in place of $\mathsf{s}(x,u)\big)$ will then hold. So, to complete the proof of implication (I) $\Longrightarrow$ (II), it remains to check condition (h"). In view of the monotonicity condition 3.27(II)(b), it is enough to show that the conjuncture of conditions $0 \leqslant x < y \leqslant \infty$, $\mathsf{s}(x+) < \mathsf{s}(x)$, and $\mathsf{s}(y) = \mathsf{s}(x+)$ effects a contradiction. Now, conditions $\mathsf{s}(y) = \mathsf{s}(x+)$ and 3.27(II)(b) imply that for all $z \in (x,y)$ one has $\mathsf{s}(z) = \mathsf{s}(y)$ and hence, by 3.31(II)(h'), $\nu\big((z,y)\big) = 0$, for $\nu := \mu_+$. So, $\nu\big((x,y)\big) = \lim_{z \downarrow x} \nu\big((z,y)\big) = 0$. Using 3.31(II)(h') again, one has $\mathsf{s}(x) = \mathsf{s}(y)$. This contradicts the assumptions $\mathsf{s}(x+) < \mathsf{s}(x)$ and $\mathsf{s}(y) = \mathsf{s}(x+)$.

$\boxed{\text{Checking (II)} \Longrightarrow \text{(I).}}$ Here it is assumed that condition (II) of Proposition 3.32 takes place. Let $\varphi(z) := \frac{z}{1+z}$ for $z \in [0,\infty)$ and $\varphi(\infty) := 1$. Then, in view of conditions 3.27(II)(a)–(c), the formulas

$\sigma\big((-\infty,0)\big):=0$ and

$$\sigma\big([0,x)\big):=\varphi\big(-\mathsf{s}(x)\big) \tag{4.48}$$

for all $x\in[0,\infty]$ uniquely determine a finite nonnegative measure σ on $\mathscr{B}(\mathbb{R})$.

Observe that $\operatorname{supp}\sigma$ does not contain isolated points. Indeed, suppose that there exists an isolated point $x\in\operatorname{supp}\sigma$. Then there exists an open set $O\subseteq\mathbb{R}$ such that $x\in O$ but $(O\setminus\{x\})\cap\operatorname{supp}\sigma=\emptyset$. It follows that $\sigma(\{x\})=\sigma(O)>0$, and so, $\varphi\big(-\mathsf{s}(x+)\big)-\varphi\big(-\mathsf{s}(x)\big)=\sigma(\{x\})>0$, whence $\mathsf{s}(x+)<\mathsf{s}(x)$. Now, for any $y\in(x,\infty]$, (h”) yields $\mathsf{s}(y)<\mathsf{s}(x+)$, so that $\sigma\big((x,y)\big)>0$ and $(x,y)\cap\operatorname{supp}\sigma\neq\emptyset$. This contradicts the assumption that x is an isolated point of $\operatorname{supp}\sigma$. Recall also that the support of any measure is a closed set.

So, by a well-known fact (see e.g. [23, Corollary 6.2], or [1, Theorem 5.3] together with [2, Problem 11.4.5(b)]), there exists a non-atomic probability measure, say ν_0, on $\mathscr{B}(\mathbb{R})$ with

$$\operatorname{supp}\nu_0=\operatorname{supp}\sigma. \tag{4.49}$$

Now one can see that condition 3.31(II)(h’) is satisfied for ν_0 in place of ν. Indeed, take any x and y such that $0\leqslant x<y\leqslant\infty$. Then (4.49), (4.48), and (h”) imply

$$\nu_0\big((x,y)\big)=0\iff\sigma\big((x,y)\big)=0\iff\mathsf{s}(x+)=\mathsf{s}(y)\implies\mathsf{s}(x+)=\mathsf{s}(x).$$

So, $\nu_0\big((x,y)\big)=0$ implies $\mathsf{s}(x)=\mathsf{s}(y)$. Vice versa, in view of the monotonicity, $\mathsf{s}(x)=\mathsf{s}(y)$ implies that s is constant on $[x,y]$, whence $\mathsf{s}(x+)=\mathsf{s}(y)$, $\sigma\big((x,y)\big)=0$, and thus $\nu_0\big((x,y)\big)=0$. This verifies condition 3.31(II)(h’) for ν_0 in place of ν.

Next, by (3.47) and (3.48), $M_{\mathsf{s}}(0)=[0,a]\times[0,1]$. Also, $\sigma\big([0,a)\big)=\varphi\big(-\mathsf{s}(a)\big)=0$, whence, by (4.49), $\nu_0\big([0,a)\big)=0$. Since ν_0 is non-atomic, $\nu_0\big([0,a]\big)=0$. Also, $\nu_0\big((-\infty,0)\big)=0$, since $\operatorname{supp}\nu_0=\operatorname{supp}\sigma\subseteq[0,\infty)$. So, $\mathsf{s}<0$ on $\operatorname{supp}\nu_0$. Thus, for any $c\in(0,\infty)$, the formula

$$\nu(A):=c\int_A\frac{\nu_0(\mathrm{d}z)}{1+z-\frac{z}{\mathsf{s}(z)}}\quad\text{for all }A\in\mathscr{B}(\mathbb{R})$$

$\big(\frac{z}{\mathsf{s}(z)}:=0$ if $\mathsf{s}(z)=-\infty\big)$ correctly defines a finite non-atomic measure on $\mathscr{B}(\mathbb{R})$, and at that the measures ν and ν_0 are absolutely continuous relative each other.

Hence, ν satisfies condition 3.31(II)(h’). Also, conditions 3.27(II)(e,f) obviously hold, as well as condition 3.31(II)(k”) – provided that c is chosen appropriately.

To complete the proof of implication (II) $\implies$ (I), it remains to check conditions 3.27(II)(i,j) – for $\mathsf{s}(x)$ in place of $\mathsf{s}(x,u)$.

Checking 3.27(II)(i). Assume that $x\in[0,\infty)$ and $\mathsf{s}(x)=-\infty$. Then, by 3.27(II)(b), $\mathsf{s}(\infty)=-\infty=\mathsf{s}(x)$. So, by the already verified condition 3.31(II)(h’), $\nu\big((x,\infty)\big)=0$. Hence, $G_\nu(x)=G_\nu(\infty)-\int_{(x,\infty)}z\,\nu(\mathrm{d}z)=G_\nu(\infty)=m_\nu$.

Checking 3.27(II)(j). Assume that $x\in(0,\infty]$ and $\mathsf{s}(x)=0$. Then, by 3.27(II)(a), $\mathsf{s}(0)=0=\mathsf{s}(x)$. So, by 3.31(II)(h’), $\nu\big((0,x)\big)=0$. Hence and because ν is non-atomic, $\nu\big((0,x]\big)=0$. It follows that $G_\nu(x)=\int_{(0,x]}z\,\nu(\mathrm{d}z)=0$.

Thus, implication (II) $\implies$ (I) is proved. It remains to note that the uniqueness of r follows by Proposition 3.30. $\square$

Proof of Proposition 3.33.

$\boxed{\text{Checking (I)} \Longrightarrow \text{(II):}}$ Assume that condition (I) takes place. Then conditions $\mathsf{s}(0) = 0$, $\nu\big((-\infty,0)\big) = 0$, $m_\nu = G_\nu(\infty) < \infty$, and 3.31(II)(k") hold by Proposition 3.31. Condition (h") is obvious, since $\nu = \mu_+$ and $\operatorname{supp}\mu = I$. So, to complete the proof of implication (I) $\Longrightarrow$ (II), it remains to check conditions (b'), (c'), and (i"). At this point we shall do more: check conditions 3.35(II)(b",c",i"',$\mathsf{r}\circ\mathsf{r}$).

Note that the function G is continuous on $[-\infty,\infty]$, since μ is non-atomic (again, we omit the subscript $_\mu$ everywhere here). Because $\operatorname{supp}\mu = I = \mathbb{R}\cap[a_-,a_+]$, the restriction (say G_+) of G to the interval $[0,a_+]$ is strictly increasing and maps $[0,a_+]$ onto $[0,m]$. Similarly, the restriction (say G_-) of G to $[a_-,0]$ is strictly decreasing and maps $[a_-,0]$ onto $[0,m]$. Hence, the function $x_+\colon[0,m]\to\mathbb{R}$ is continuous and strictly increasing, as the inverse to the continuous and strictly increasing function G_+, and x_+ maps $[0,m]$ onto $[0,a_+]$. Similarly, the function $x_-\colon[0,m]\to\mathbb{R}$ is continuous and strictly decreasing, and it maps $[0,m]$ onto $[a_-,0]$. Now conditions 3.35(II)(b",c") follow by (2.6), since $\mathsf{r}(x) = x_+\big(G(x)\big)$ for $x\in[0,\infty]$ and $\mathsf{r}(x) = x_-\big(G(x)\big)$ for $x\in[-\infty,0]$.

Take now any $x\in[a_-,0]$ and let $y := \mathsf{r}(x) = x_+\big(G(x)\big) = G_+^{-1}\big(G_-(x)\big)\in[0,a_+]$. Then $G_+(y) = G_-(x)$ and $\mathsf{r}\big(\mathsf{r}(x)\big) = \mathsf{r}(y) = x_-\big(G(y)\big) = G_-^{-1}\big(G_+(y)\big) = G_-^{-1}\big(G_-(x)\big) = x$. Similarly, $\mathsf{r}\big(\mathsf{r}(x)\big) = x$ for all $x\in[0,a_+]$. This proves condition 3.35(II)($\mathsf{r}\circ\mathsf{r}$) as well.

Further, take any $x\in[-\infty,a_-]$. Then $G(x) = G(a_-) = m$, since $\operatorname{supp}\mu = I = \mathbb{R}\cap[a_-,a_+]$ does not intersect with $[-\infty,a_-)$. Hence, $\mathsf{r}(x) = x_+\big(G(x)\big) = G_+^{-1}(m) = a_+$. Similarly, $\mathsf{r}(x) = a_-$ for all $x\in[a_+,\infty]$. This proves condition 3.35(II)(i"') and thus completes the entire proof of implication (I) $\Longrightarrow$ (II). Moreover, we have shown that (I) implies conditions 3.35(b",c",i"',$\mathsf{r}\circ\mathsf{r}$).

$\boxed{\text{Checking (II)} \Longrightarrow \text{(I):}}$ Assume that condition (II) takes place. Then, by Proposition 3.31, it suffices to check conditions 3.27(II)(b,c,i,j), 3.31(II)(h'), and $\operatorname{supp}\mu = I$. Conditions 3.27(II)(b) and 3.27(II)(c) follow immediately from 3.33(II)(b',i") and, respectively, 3.33(II)(c',i").

$\boxed{\text{Checking 3.27(II)(i):}}$ Take any $x\in[0,\infty)$ such that $G_\nu(x) < m_\nu$. Then $x\in[0,a_+)$; indeed, in view of 3.33(II)(h") – and because ν is non-atomic and hence $\nu(\{a_+\}) = 0$, one has $G_\nu(y) = G_\nu(\infty) = m_\nu$ for all $y\in[a_+,\infty]$. So, by 3.33(II)(b'), $\mathsf{s}(x) > \mathsf{s}(a_+) \geqslant -\infty$, whence $\mathsf{s}(x) > -\infty$.

$\boxed{\text{Checking 3.27(II)(j):}}$ Take any $x\in(0,\infty]$. Then, by 3.33(II)(i",b'), $\mathsf{s}(x) = \mathsf{s}(x\wedge a_+) < \mathsf{s}(0) = 0$, whence $\mathsf{s}(x) < 0$.

$\boxed{\text{Checking 3.31(II)(h'):}}$ Take any x and y such that $0\leqslant x < y\leqslant\infty$. We have to check the equivalence $\nu\big((x,y)\big) = 0 \iff \mathsf{s}(x) = \mathsf{s}(y)$.

Assume here first that $\nu\big((x,y)\big) = 0$. Then $(x,y)\cap[0,a_+] = \emptyset$, by 3.33(II)(h"). So, $a_+\leqslant x < y\leqslant\infty$, whence, by by 3.33(II)(i"), $\mathsf{s}(x) = a_- = \mathsf{s}(y)$.

Vice versa, assume that $\mathsf{s}(x) = \mathsf{s}(y)$. Then $a_+\leqslant x < y\leqslant\infty$ $\big($otherwise, one would have $0\leqslant x < a_+$, and so, by 3.33(II)(b',i"), $\mathsf{s}(x) > \mathsf{s}(y\wedge a_+) \geqslant \mathsf{s}(y)$, a contradiction$\big)$. Hence, $(x,y)\subseteq(a_+,\infty)\subseteq\mathbb{R}\setminus\operatorname{supp}\nu$, by 3.33(II)(h"). So, $\nu\big((x,y)\big) = 0$.

$\boxed{\text{Checking } \operatorname{supp}\mu = I\text{:}}$ First, by 3.33(II)(h"), $G(a_+) = G(\infty) = m$. So, by 3.33(II)(i"), $a_- = \mathsf{s}(a_+) = x_-\big(G(a_+)\big) = x_-(m)$. Hence, $m = G(-\infty)\geqslant G(a_-)\geqslant m$, by (3.2). So, $G(-\infty) = m = G(a_-)$, whence $\mu\big((-\infty,a_-)\big) = 0$. Therefore, $(-\infty,a_-)\cap\operatorname{supp}\mu = \emptyset$.

Next, take any z_1 and z_2 such that $a_- < z_1 < z_2 < 0$. By conditions $\mathsf{s}(0) = 0$ and 3.33(II)(b',c',i"), the function s continuously decreases on the interval $[0,a_+]$ from 0 to a_-. So, there exist x_1 and

x_2 such that $z_1 = \mathsf{s}(x_1)$, $z_2 = \mathsf{s}(x_2)$, and $0 < x_2 < x_1 < a_+$. Then, by (4.44), $\mathsf{r}(z_1) = \mathsf{r}(\mathsf{s}(x_1)) = x_1$ and $\mathsf{r}(z_2) = \mathsf{r}(\mathsf{s}(x_2)) = x_2$, whence $\mathsf{r}(z_1) \neq \mathsf{r}(z_2)$. Now, by the natural left "mirror" analogue of 3.31(II)(h'), $\mu((z_1, z_2)) > 0$. So, $\operatorname{supp}\mu \supseteq \mathbb{R} \cap [a_-, 0]$. Recalling now that $(-\infty, a_-) \cap \operatorname{supp}\mu = \emptyset$, one has $(-\infty, 0] \cap \operatorname{supp}\mu = \mathbb{R} \cap [a_-, 0]$. Also, $[0, \infty) \cap \operatorname{supp}\mu = \operatorname{supp}\nu = \mathbb{R} \cap [0, a_+]$. Thus, indeed $\operatorname{supp}\mu = \mathbb{R} \cap [a_-, a_+] = I$.

This completes the proof of implication (II) $\Longrightarrow$ (I). It remains to note that the uniqueness of μ follows immediately by Proposition 3.27. $\square$

Proof of Proposition 3.34. The implication (I) $\Longrightarrow$ (II) follows immediately from Proposition 3.33.

The uniqueness of r follows immediately from Proposition 3.30 or Proposition 3.32.

Finally, implication (II) $\Longrightarrow$ (I) follows by Proposition 3.32, since conditions $\mathsf{s}(0) = 0$ and 3.33(II)(b',c',i") imply 3.27(II)(a–c) and 3.32(II)(h"). $\square$

Proof of Proposition 3.35. Implication (I) $\Longrightarrow$ (II) of Proposition 3.35 was already proved in the proof of implication (I) $\Longrightarrow$ (II) of Proposition 3.33.

Implication (II) $\Longrightarrow$ (I) of Proposition 3.35 follows immediately by Proposition 3.34. $\square$

Proof of Proposition 3.36. $\boxed{\text{Checking (I)} \Longrightarrow \text{(II):}}$ Assume that condition (I) takes place. Then, by Proposition 3.35, condition 3.35(II) holds. Since $\operatorname{supp}\mu = I$ and in an open neighborhood (say O) of 0 measure μ has a continuous strictly positive density (say f), function G is continuously differentiable in O, with $G'(x) = x f(x)$ for all $x \in O$. Also, by the continuity of r, one has $\mathsf{r}(x) \in O$ for all x in some other open neighborhood (say O_1) of 0 such that $O_1 \subseteq O$. Next, by (2.6) and (say) (3.4), (3.6), one has

$$G(\mathsf{r}(x)) = G(x) \quad \text{for all } x \in [-\infty, \infty]; \tag{4.50}$$

once again, the subscript $_\mu$ is omitted. So, by the inverse function theorem, r is differentiable in $O_1 \setminus \{0\}$ and, for all $x \in O_1 \setminus \{0\}$,

$$\mathsf{r}'(x) = \frac{G'(x)}{G'(\mathsf{r}(x))} = \frac{x f(x)}{\mathsf{r}(x) f(\mathsf{r}(x))} \sim \frac{x}{\mathsf{r}(x)} \tag{4.51}$$

as $x \to 0$. In particular, it follows by (4.51) that r' is continuous in $O_1 \setminus \{0\}$. It also follows that $\mathsf{r}(x)^2 = \int_0^x 2\mathsf{r}(z)\,\mathsf{r}'(z)\,dz \sim \int_0^x 2z\,dz = x^2$, $\mathsf{r}(x) \sim -x$, $\mathsf{r}'(0) = -1$, and, again by (4.51), $\mathsf{r}'(x) \to -1$ as $x \to 0$, so that r' is continuous at 0 as well. This completes the proof of implication (I) $\Longrightarrow$ (II).

$\boxed{\text{Checking (II)} \Longrightarrow \text{(I):}}$ Assume that condition (II) takes place. Let $\mathsf{s} := \mathsf{r}_+$. Then one can easily construct a measure ν on $\mathscr{B}(\mathbb{R})$ with a density $g := \frac{d\nu}{dx}$ that is continuous and strictly positive on $[0, a_+)$ and such that condition 3.33(II) holds. Then condition 3.33(I) holds as well, so that there exists a non-atomic zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$ such that $\operatorname{supp}\mu = I$, $\mu_+ = \nu$, and the reciprocating function $(\mathsf{r}_\mu)_+ = \mathsf{s}$. Moreover, by the uniqueness part of Proposition 3.34, $\mathsf{r}_\mu = \mathsf{r}$. Therefore, identity (4.50) holds with $G = G_\mu$. So, for all x in a left neighborhood (l.n.) of 0 there exists the derivative

$$G'(x) = G'(\mathsf{r}(x))\,\mathsf{r}'(x) = \mathsf{r}(x)\,g(\mathsf{r}(x))\,\mathsf{r}'(x) \sim x\,g(0) \tag{4.52}$$

as $x \uparrow 0$, because $\mathsf{r}'(0-) = \mathsf{r}'(0) = -1$ and $\mathsf{r}(x) \sim -x$ as $x \uparrow 0$. On the other hand, (4.52) implies that G' is strictly negative and continuous in a l.n. (say O_-) of 0. So, for all $A \in \mathscr{B}(O_-)$,

$$\mu(A) = \int_A \frac{\mathrm{d}G(x)}{x} = \int_A \frac{G'(x)\,\mathrm{d}x}{x} = \int_A f(x)\,\mathrm{d}x,$$

where $f(x) := \frac{G'(x)}{x} = \frac{\mathsf{r}(x)}{x}\, g\big(\mathsf{r}(x)\big)\, \mathsf{r}'(x)$ is continuous on O_- and, by (4.52), $f(0-) = g(0)$. Gluing the functions f and g together, one sees that indeed the probability measure μ has a continuous strictly positive density a neighborhood of 0. $\square$

Proof of Proposition 3.38.

$\boxed{\text{Checking (I):}}$ By conditions $\mathsf{r}(0) = 0$ and 3.35(II)(b",c",i"'), for x in the interval $[0, a_+)$ the width $\mathsf{w}(x) = |x - \mathsf{r}(x)|$ equals $x - \mathsf{r}(x)$ and hence continuously and strictly increases from 0 to $a_+ - a_-$; therefore, the restriction of function w to the interval $[0, a_+)$ has a unique inverse, say w_+^{-1}, which continuously increases on $[0, a_+ - a_-)$ from 0 to a_+. Similarly, for y in the interval $(a_-, 0]$, $\mathsf{w}(y)$ equals $\mathsf{r}(y) - y$ and hence continuously and strictly decreases from $a_+ - a_-$ to 0, so that the restriction of function w to the interval $(a_-, 0]$ has a unique inverse, say w_-^{-1}, which continuously decreases on $[0, a_+ - a_-)$ from 0 to a_-. Thus, condition (3.50) will hold for all $x \in [0, a_+)$ iff $\mathsf{a}(w) = \mathsf{a}_+(w) := \big(x + \mathsf{r}(x)\big)\big|_{x = \mathsf{w}_+^{-1}(w)}$ for all $w \in [0, a_+ - a_-)$; similarly, (3.50) will hold for all $y \in (a_-, 0]$ iff $\mathsf{a}(w) = \mathsf{a}_-(w) := \big(y + \mathsf{r}(y)\big)\big|_{y = \mathsf{w}_-^{-1}(w)}$ for all $w \in [0, a_+ - a_-)$. So, to prove the existence and uniqueness of a function a satisfying condition (3.50) for all $w \in [0, a_+ - a_-)$, it suffices to show that $\mathsf{a}_+(w) = \mathsf{a}_-(w)$ for all $w \in [0, a_+ - a_-)$, that is, to verify the implication

$$x - \mathsf{r}(x) = \mathsf{r}(y) - y \implies y + \mathsf{r}(y) = x + \mathsf{r}(x) \tag{4.53}$$

whenever $a_- < y \leqslant 0 \leqslant x < a_+$. Fix any such x and y. Again by conditions $\mathsf{r}(0) = 0$ and 3.35(II)(b",c",i"'), r maps interval $(a_-, 0]$ onto $[0, a_+)$. Hence, there exists some $\tilde{y} \in (a_-, 0]$ such that $x = \mathsf{r}(\tilde{y})$, so that, by condition 3.35(II)($\mathsf{r} \circ \mathsf{r}$), $\mathsf{r}(x) = \tilde{y}$. It follows that $\mathsf{r}(y) - y = x - \mathsf{r}(x) = \mathsf{r}(\tilde{y}) - \tilde{y}$, and so, $y = \tilde{y}$, since $\mathsf{r}(y) - y$ strictly decreases in $y \in \mathbb{R}$. Therefore, $\mathsf{r}(y) + y = \mathsf{r}(\tilde{y}) + \tilde{y} = x + \mathsf{r}(x)$, so that implication (4.53) is verified.

Next, let us check the strict Lip(1) condition, which is easy to see to be equivalent to the condition that the functions ξ and ρ (defined by (3.52)) are strictly increasing on $[0, a_+ - a_-)$.

For each $w \in [0, a_+ - a_-)$, $x := \mathsf{w}_+^{-1}(w)$, and $y := \mathsf{w}_-^{-1}(w)$, one has $w = \mathsf{w}(x) = \mathsf{w}(y)$, $x \in [0, a_+)$, and $y \in (a_-, 0]$, and so, by (3.52) and (3.50),

$$\xi(w) = \tfrac{1}{2}(w + \mathsf{a}(w)) = \tfrac{1}{2}(x - \mathsf{r}(x) + x + \mathsf{r}(x)) = \mathsf{w}_+^{-1}(w) \quad \text{and} \tag{4.54}$$

$$\xi(w) = \tfrac{1}{2}(w + \mathsf{a}(w)) = \tfrac{1}{2}(\mathsf{r}(y) - y + y + \mathsf{r}(y)) = \mathsf{r}\big(\mathsf{w}_-^{-1}(w)\big); \tag{4.55}$$

either of these two lines shows that ξ is continuously and strictly increasing on $[0, a_+ - a_-)$, from 0 to a_+; similarly,

$$\rho(w) = \tfrac{1}{2}(w - \mathsf{a}(w)) = \tfrac{1}{2}(x - \mathsf{r}(x) - x - \mathsf{r}(x)) = -\mathsf{r}\big(\mathsf{w}_+^{-1}(w)\big) \quad \text{and} \tag{4.56}$$

$$\rho(w) = \tfrac{1}{2}(w - \mathsf{a}(w)) = \tfrac{1}{2}(\mathsf{r}(y) - y - y - \mathsf{r}(y)) = -\mathsf{w}_-^{-1}(w); \tag{4.57}$$

either of the last two lines shows that ρ is continuously and strictly increasing on $[0, a_+ - a_-)$, from 0 to $-a_-$. It also follows that $\mathsf{a}(w) = \xi(w) - \rho(w) \to a_+ + a_-$ as $w \uparrow a_+ - a_-$. Thus, statement (I) is verified.

$\boxed{\text{Checking (II):}}$ As noted in the above proof of statement (I), the strict Lip(1) condition on a is equivalent to the condition that the functions ξ and ρ be strictly increasing on $[0, a_+ - a_-)$; these functions are also continuous, in view of definition (3.52), since the function a is Lipschitz and hence continuous. Therefore, the functions ξ and ρ are strictly and continuously increasing on $[0, a_+ - a_-)$, from 0 to a_+ and $-a_-$, respectively. So, the functions ξ and ρ have strictly and continuously increasing inverses ξ^{-1} and ρ^{-1}, which map $[0, a_+)$ and $[0, -a_-)$, respectively, onto $[0, a_+ - a_-)$.

Thus, one can use formula (3.51) *to define a* function r on the interval (a_-, a_+). Let us then extend this definition to the entire interval $[-\infty, \infty]$ by imposing condition 3.35(II)(i'''). Then one can see that the function r satisfies condition 3.35(II). So, Proposition 3.35 implies that r is the reciprocating function of a nonatomic zero-mean probability measure μ on $\mathscr{B}(\mathbb{R})$ with $\operatorname{supp}\mu = I = \mathbb{R} \cap [a_-, a_+]$.

Let us now verify (3.50). Take any $x \in [0, a_+)$. Then $x = \xi(w)$ for some $w \in [0, a_+ - a_-)$, whence $\xi^{-1}(x) = w$ and, in view of (3.51),

$$|x - \mathsf{r}(x)| = x - \mathsf{r}(x) = x + \rho\big(\xi^{-1}(x)\big) = \xi(w) + \rho(w) = w,$$
$$x + \mathsf{r}(x) = x - \rho\big(\xi^{-1}(x)\big) = \xi(w) - \rho(w) = a(w),$$

so that (3.50) holds for $x \in [0, a_+)$. Similarly, (3.50) holds for $x \in (a_-, 0]$. So, (3.50) is verified.

To complete the proof of (II), it remains to check the uniqueness of r given a and (3.50). That is, we have to show that the value of $\mathsf{r}(x)$ is uniquely determined for each $x \in (a_-, a_+)$. In fact, we shall show that, moreover, relations (3.51) must necessarily hold. Toward that end, observe that, as shown in the above proof of statement (I), condition (3.50) implies (4.54)–(4.57).

Now, take any $x \in [0, a_+)$. Then $x = \mathsf{w}_+^{-1}(w)$ for some $w \in [0, a_+ - a_-)$, whence, by (4.54), $\xi(w) = x$ and hence $\xi^{-1}(x) = w$; so, by (4.56), $-\rho\big(\xi^{-1}(x)\big) = -\rho(w) = \mathsf{r}\big(\mathsf{w}_+^{-1}(w)\big) = \mathsf{r}(x)$, which proves the first case in (3.51).

Similarly, take any $y \in (a_-, 0]$. Then $y = \mathsf{w}_-^{-1}(w)$ for some $w \in [0, a_+ - a_-)$, whence, by (4.57), $\rho(w) = -y$ and hence $\rho^{-1}(-y) = w$. So, by (4.55), $\xi\big(\rho^{-1}(-y)\big) = \xi(w) = \mathsf{r}\big(\mathsf{w}_-^{-1}(w)\big) = \mathsf{r}(y)$, which proves the second case in (3.51).

$\boxed{\text{Checking (III), the “if” part:}}$ Here, assume that the asymmetry pattern function a is continuously differentiable in an open r.n. of 0 and $\mathsf{a}'(0+) = 0$. Then in such a r.n. the functions ξ and ρ are continuously differentiable, $\xi' = \frac{1}{2}(1 + \mathsf{a}')$, $\xi'(0+) = \frac{1}{2}$, $\rho' = \frac{1}{2}(1 - \mathsf{a}')$, $\rho'(0+) = \frac{1}{2}$. Recall also that the functions ξ and ρ are continuously increasing on $[0, a_+ - a_-)$, so that the inverse functions ξ^{-1} and ρ^{-1} are continuously differentiable in a r.n. of 0. Hence, by (3.51), there is some $\varepsilon > 0$ such that

$$x \in (0, \varepsilon) \implies \mathsf{r}'(x) = \frac{-2\rho'\big(\xi^{-1}(x)\big)}{1 + \mathsf{a}'\big(\xi^{-1}(x)\big)} \underset{x \downarrow 0}{\longrightarrow} -1,$$
$$x \in (-\varepsilon, 0) \implies \mathsf{r}'(x) = \frac{-2\xi'\big(\rho^{-1}(-x)\big)}{1 - \mathsf{a}'\big(\rho^{-1}(-x)\big)} \underset{x \uparrow 0}{\longrightarrow} -1,$$

which shows, in view of the mean value theorem, that indeed the corresponding reciprocating function r is continuously differentiable in a neighborhood of 0.

$\boxed{\text{Checking (III), the “only if” part:}}$ Here, assume that a reciprocating function r such as in Proposition 3.35 is continuously differentiable in a neighborhood of 0. Then, by Proposition 3.36,

$\mathsf{r}'(0) = -1$. By (3.50), $\mathsf{a} \circ \mathsf{w}_+ = \alpha$, where $\mathsf{w}_+ \colon [0, a_+) \to [0, a_+ - a_-)$, $\alpha \colon [0, a_+) \to \mathbb{R}$, $\mathsf{w}_+(x) \equiv x - \mathsf{r}(x)$ and $\alpha(x) \equiv x + \mathsf{r}(x)$. At that, the function w_+ is continuously differentiable in some r.n., say $(0, \varepsilon)$, of 0, with $\mathsf{w}_+'(x) = 1 - \mathsf{r}'(x)$ for all $x \in (0, \varepsilon)$, so that $\mathsf{w}_+'(0+) = 2 \neq 0$. So, for all w in some open r.n. of 0, one has $\mathsf{a}(w) = \alpha\big(\mathsf{w}_+^{-1}(w)\big)$ and hence

$$\mathsf{a}'(w) = \frac{1 + \mathsf{r}'\big(\mathsf{w}_+^{-1}(w)\big)}{\mathsf{w}_+'\big(\mathsf{w}_+^{-1}(w)\big)} \underset{w \downarrow 0}{\longrightarrow} 0,$$

so that indeed the function a is continuously differentiable in an open right neighborhood of 0 and $\mathsf{a}'(0+) = 0$. ☐

4.2 Proofs of the main results

Proof of Theorem 2.2. This theorem is a special case of Proposition 3.13. ☐

Proof of Theorem 2.4. For $j = 1, \dots, n+1$, introduce

$$g_j(p_1, \dots, p_n) := \mathsf{E}\, g(p_1, \dots, p_{j-1}, X_{j;p_j}, R_{j;p_j}, \dots, X_{n;p_n} R_{n;p_n}) \tag{4.58}$$

and

$$\mathscr{I}_j := \int_{(\mathbb{R}\times[0,1])^n} g_j(p_1, \dots, p_n)\, \mathrm{d}p_1 \dots \mathrm{d}p_n.$$

Then, for all $j = 1, \dots, n$,

$$\mathscr{I}_{j+1} = \int_{(\mathbb{R}\times[0,1])^{n-1}} \mathscr{E}_j\, \mathrm{d}p_1 \dots \mathrm{d}p_{j-1}\, \mathrm{d}p_{j+1} \dots \mathrm{d}p_n,$$

where

$$\begin{aligned}
\mathscr{E}_j &:= \mathsf{E}\, g_{j+1}(p_1, \dots, p_{j-1}, X_j, R_j, p_{j+1}, \dots, p_n) \\
&= \int_{\mathbb{R}\times[0,1]} \mathsf{E}\, g_{j+1}(p_1, \dots, p_{j-1}, X_{j;p_j}, R_{j;p_j}, p_{j+1}, \dots, p_n)\, \mathsf{P}(X_j \in \mathrm{d}x_j)\, \mathrm{d}u_j \\
&= \int_{\mathbb{R}\times\mathbb{R}\times[0,1]} g_{j+1}(p_1, \dots, p_{j-1}, x_{j;p_j}, r_{j;p_j}, p_{j+1}, \dots, p_n) \\
&\qquad\qquad \times \mathsf{P}\left((X_{j;p_j}, R_{j;p_j}) \in \mathrm{d}x_{j;p_j} \times \mathrm{d}r_{j;p_j}\right) \mathsf{P}(X_j \in \mathrm{d}x_j)\mathrm{d}u_j \\
&= \int_{\mathbb{R}\times\mathbb{R}\times[0,1]} \mathsf{E}\, g(p_1, \dots, p_{j-1}, x_{j;p_j}, r_{j;p_j}, X_{j+1;p_{j+1}}, R_{j+1;p_{j+1}}, \dots, X_{n;p_n}, R_{n;p_n}) \\
&\qquad\qquad \times \mathsf{P}\left((X_{j;p_j}, R_{j;p_j}) \in \mathrm{d}x_{j;p_j} \times \mathrm{d}r_{j;p_j}\right) \mathsf{P}(X_j \in \mathrm{d}x_j)\mathrm{d}u_j \\
&= \int_{\mathbb{R}\times[0,1]} \mathsf{E}\, g(p_1, \dots, p_{j-1}, X_{j;p_j}, R_{j;p_j}, \dots, X_{n;p_n}, R_{n;p_n})\, \mathsf{P}(X_j \in \mathrm{d}x_j)\, \mathrm{d}u_j \\
&= \int_{\mathbb{R}\times[0,1]} g_j(p_1, \dots, p_n)\, \mathsf{P}(X_j \in \mathrm{d}x_j)\, \mathrm{d}u_j = \int_{\mathbb{R}\times[0,1]} g_j(p_1, \dots, p_n)\, \mathrm{d}p_j;
\end{aligned}$$

the second of these 7 equalities follows by (2.12), and the fourth and sixth ones by (4.58). Now it follows that $\mathscr{I}_{j+1}=\mathscr{I}_j$, for all $j=1,\dots,n$. This finally implies $\mathscr{I}_{n+1}=\mathscr{I}_1$, so that

$$\begin{aligned}&\mathsf{E}\,g(X_1,R_1,\dots,X_n,R_n)\\&\qquad=\mathscr{I}_{n+1}=\mathscr{I}_1=\int_{(\mathbb{R}\times[0,1])^n}\mathsf{E}\,g(X_{1;p_1}R_{1;p_1},\dots,X_{n;p_n},R_{n;p_n})\,\mathrm{d}p_1\dots\mathrm{d}p_n.\end{aligned}$$

□

Proof of Corollary 2.5. Take any function $f\in\mathscr{H}_+^5$ and, for any $x_1,r_1\dots,\ x_n,r_n$ in $\mathbb{R}$, let

$$g_f(x_1,r_1\dots,x_n,r_n):=\begin{cases}f\left(\dfrac{x_1+\cdots+x_n}{\frac12\sqrt{w_1^2+\cdots+w_n^2}}\right)&\text{if } w_1^2+\cdots+w_n^2\neq0,\\ f(0)&\text{otherwise,}\end{cases}$$

where $w_i:=|x_i-r_i|$. Then, by Theorem 2.4 and [26, Theorem 2.1],

$$\begin{aligned}\mathsf{E}\,f(S_W)&=\mathsf{E}\,g_f(X_1,R_1,\dots,X_n,R_n)\\&\leqslant\sup\big\{\,\mathsf{E}\,g_f(X_{x_1,r_1},R_{x_1,r_1},\dots,X_{x_n,r_n},R_{x_n,r_n})\colon(x_1,r_1\dots,x_n,r_n)\in\mathbb{R}^{2n},\\&\qquad\qquad x_jr_j\leqslant0\ \forall j\big\}\\&=\sup\left\{\,\mathsf{E}\,f\left(\frac{X_{x_1,r_1}+\cdots+X_{x_n,r_n}}{\frac12\sqrt{(x_1-r_1)^2+\cdots+(x_n-r_n)^2}}\right)\colon(x_1,r_1\dots,x_n,r_n)\in\mathbb{R}^{2n},\right.\\&\qquad\qquad\left.x_jr_j\leqslant0\ \forall j\right\}\\&\leqslant\mathsf{E}\,f(Z),\end{aligned}$$

which proves (2.14). Now (2.15) follows by [26, Corollary 2.2]. □

Proof of Corollary 2.6. This proof is similar to that of Corollary 2.5, using [28, Theorem 4 and Corollary 3] instead of [26, Theorem 2.1 and Corollary 2.2]. Here we only would like to provide some details concerning condition (2.16), in its relation with condition [28, (20)]. Namely, we shall show that (2.16) implies that

$$\frac{\mathsf{r}_i(y,u)}{|y|}\leqslant\gamma:=\frac{1-p}{p}\quad\text{for all } y<0,\ u\in[0,1],\tag{4.59}$$

and $i\in\{1,\dots,n\}$, which will allow one to immediately apply the mentioned results of [28].

Fix any $i\in\{1,\dots,n\}$ and write, for brevity, X and r for X_i and r_i, respectively. Then, by (2.16) and Fubini's theorem, one has $\mathsf{P}(X\in A)=1$, where $A:=\{x>0\colon \mathrm{mes}(B_x)=1\}$, mes denotes the Lebesgue measure, and $B_x:=\{u\in[0,1]\colon\frac{x}{|\mathsf{r}(x,u)|}\leqslant\gamma\}$. Since B_x is a closed interval, one has $B_x=[0,1]$ and hence $\frac{x}{|\mathsf{r}(x,0)|}\leqslant\gamma$ for all $x\in A$.

Take now any $y<0$ and let $x_y:=\mathsf{r}(y,1)$ and $h:=\tilde G(y,1)=G(y)$. Then $x_y=x_+(h)$, $G(x_y-)\leqslant h$ $\big($by (3.4)$\big)$, and $0\geqslant\mathsf{r}(x_y,0)=x_-(G(x_y-))\geqslant x_-(h)\geqslant y$ $\big($by property (i) of Proposition 3.1 and (3.2)$\big)$, so that $|\mathsf{r}(x_y,0)|\leqslant|y|$ and

$$\frac{\mathsf{r}(y,u)}{|y|}\leqslant\frac{x_y}{|y|}\leqslant\frac{x_y}{|\mathsf{r}(x_y,0)|}\quad\text{for all } u\in[0,1].\tag{4.60}$$

Moreover, w.l.o.g. $x_y > 0$ $\big($otherwise, $\mathsf{r}(y,1) = x_y = 0$ and hence $\mathsf{r}(y,u) = 0$ for all $u \in [0,1]$, so that (4.59) is trivial$\big)$. On the other hand, for each $x \in [0, x_y)$, by (3.3) and (3.1) one has $G(x) < h \leqslant G(x_y)$ and hence $\mathsf{P}(X \in (x, x_y]) > 0$, so that $(x, x_y] \cap A \neq \emptyset$ (since $\mathsf{P}(X \in A) = 1$). Therefore, there exists a non-decreasing sequence (x_n) in A such that $x_n \uparrow x_y$. So, $\gamma_n := \frac{x_n}{|\mathsf{r}(x_n,0)|} \leqslant \gamma$ for all n and, in view of property (iv) of Proposition 3.1, $\frac{x_y}{|\mathsf{r}(x_y,0)|} = \lim_n \gamma_n \leqslant \gamma$, since $\mathsf{r}(x,0) = x_-(G(x-))$ for all $x > 0$ and the function $(0,\infty) \ni x \mapsto G(x-)$ is non-decreasing and left-continuous.
Thus, $\frac{x_y}{|\mathsf{r}(x_y,0)|} \leqslant \gamma$, whence, by (4.60), inequality(4.59) follows. $\square$

References

[1] Abramovich, Y. A. and Wickstead, A. W. (1993). Remarkable classes of unital AM-spaces. *J. Math. Anal. Appl.* **180**, 398–411. MR1251867

[2] Abramovich, Y. A.; Aliprantis, C. D. (2002). Problems in operator theory. American Mathematical Society, Providence, RI. MR1921783

[3] Aizenman, M., Germinet, F., Klein, A. and Warzel, S. (2009). On Bernoulli Decompositions for Random Variables, Concentration Bounds, and Spectral Localization. *Probab. Theory Related Fields* **143** 219–238. MR2449128

[4] Bartlett, M. S. (1935). The effect of non-normality on the t distribution. *Proc. Camb. Phil. Soc.* **31**, 223–231.

[5] Bentkus, V. (2004) On Hoeffding's inequalities. *Ann. Probab.* **32**, 1650–1673. MR2060313

[6] Bentkus, V. and Juškevičius, T. (2008) Bounds for tail probabilities of martingales using skewness and kurtosis. *Lithuanian Math. J.* **48**, 30–37. MR2398168

[7] Cambanis, S.; Simons, G.; Stout, W. (1976). Inequalities for $Ek(X,Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **36**, 285–294. MR0420778

[8] Eaton, M. L. (1970). A note on symmetric Bernoulli random variables. *Ann. Math. Statist.* **41**, 1223–1226. MR268930

[9] Eaton, M. L. (1974). A probability inequality for linear combinations of bounded random variables. *Ann. Statist.* **2**, 609–614.

[10] Eaton, M. L. and Efron, B. (1970). Hotelling's T^2 test under symmetry conditions. *J. Amer. Statist. Assoc.* **65**, 702–711. MR0269021

[11] Efron, B. (1969). Student's t test under symmetry conditions. *J. Amer. Statist. Assoc.* **64**, 1278–1302. MR0251826

[12] Gangbo, W. (1999). The Monge mass transfer problem and its applications. *Monge Ampère equation: applications to geometry and optimization* (Deerfield Beach, FL, 1997), 79–104, Contemp. Math., 226, Amer. Math. Soc., Providence, RI. MR1660743

[13] Hall, P. (1992). The bootstrap and Edgeworth expansion. Springer Series in Statistics. Springer-Verlag, New York. MR1145237

[14] HALL, P. AND WANG, Q. (2004). Exact convergence rate and leading term in central limit theorem for Student's t statistic. *Ann. Probab.* **32**, 1419–1437. MR2060303

[15] HOAGLIN, D. C. (1985). Summarizing shape numerically: The g- and h-distributions. In *Exploring Data Tables, Trends, and Shapes* (D. C. Hoaglin, F. Mosteller and J. W. Tukey, eds.) 461–514. Wiley, New York.

[16] HOEFFDING, W. (1940). Masstabinvariante korrelationstheorie. *Schr. Math. Inst. Univ. Berlin.* **5**, 179–233.

[17] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58**, 13–30. MR144363

[18] HOEFFDING, W. (1994). The collected works of Wassily Hoeffding. Springer, New York. MR1307621

[19] KAFADAR, K. (2003). John Tukey and robustness. *Statist. Sci.* **18**, 319–331. MR2056573

[20] KALLENBERG, O. (2002). Foundations of modern probability. Second edition. Springer-Verlag, New York. MR1876169

[21] LOGAN, B. F.; MALLOWS, C. L.; RICE, S. O.; SHEPP, L. A. (1973). Limit distributions of self-normalized sums. *Ann. Probab.* **1**, 788–809. MR0362449

[22] OATES, D. K. (1971). A non-compact Krein-Milman theorem. *Pacific J. Math.* **36**, 781–785. MR0288552

[23] PARTHASARATHY, K. R.; RANGA RAO, R.; VARADHAN, S. R. S. (1962). On the category of indecomposable distributions on topological groups. *Trans. Amer. Math. Soc.* **102**, 200–217. MR0153041

[24] PINELIS, I. (1994). Extremal probabilistic problems and Hotelling's T^2 test under a symmetry condition. *Ann. Statist.* **22**, 1, 357–368. MR1272088

[25] PINELIS, I. (2006). Binomial upper bounds on generalized moments and tail probabilities of (super)martingales with differences bounded from above. *IMS Lecture Notes Monograph Series, Vol. 51, 33–52,* `http://arxiv.org/abs/math/0512301`. MR2387759

[26] PINELIS, I. (2006). On normal domination of (super)martingales. *Electronic Journal of Probability,* **11**, Paper 39, 1049-1070, `http://www.math.washington.edu/~ejpecp/include/getdoc.php?id=3724&article=1648&mode=pdf`. MR2268536

[27] PINELIS, I. (2006). Student's t-test without symmetry conditions. arXiv:math/0606160v1 [math.ST], `http://arxiv.org/abs/math/0606160`.

[28] PINELIS, I. (2007). Exact inequalities for sums of asymmetric random variables, with applications. *Probability Theory and Related Fields,* **139**, 605–635, `http://www.springerlink.com/content/m002830hk5x43252/`. MR2322709

[29] RACHEV, S. T. AND RÜSCHENDORF, L. (1998). Mass transportation problems. Vol. I. Theory. Springer-Verlag, New York. MR1619170

[30] RATCLIFFE, J. F. (1968). The effect on the t-distribution of non-normality in the sampled population. *Appl. Statist.* **17**, 42–48.

[31] SHAO, Q.-M. (1997). Self-normalized large deviations. *Ann. Probab.* **25**, 285–328. MR1428510

[32] SKOROKHOD, A. V. (1965). Studies in the theory of random processes. Addison-Wesley, Reading, Mass. MR0185620

[33] TCHEN, A. H. (1980). Inequalities for distributions with given marginals. *Ann. Probab.* **8**, 814–827. MR0577318

[34] TUKEY, J. W. (1948). Some elementary problems of importance to small sample practice. *Human Biol.* **20**, 205–214.

[35] TUKEY, J. W. (1957). On the comparative anatomy of transformations. *Ann. Math. Statist.* **28**, 602–632. MR0091546