# SPITZER'S FORMULA: A SHORT PROOF

J. G. WENDEL

1. **Introduction.** Let $\{X_n\}$ be a sequence of mutually independent random variables with the same distribution, and let $\{S_n\}$ be the usual sequence of partial sums. In studying problems of first passage, ruin and the like, it is often useful to consider random variables $R_n = \max(0, S_1, S_2, \cdots, S_n)$.

Spitzer [1] has obtained a beautiful formula from which one can (in principle at least) calculate the joint distribution of any pair $(R_n, S_n)$ knowing the *individual* distributions of the first $n$ partial sums; he has in addition given several important applications. His formula takes the form of a relation between generating functions for certain characteristic functions; specifically, writing $a^{\pm}$ for $(|a| \pm a)/2$ and putting $\phi_n(\alpha, \beta) = E(\exp i(\alpha R_n + \beta(R_n - S_n)),\ u_n(\alpha) = E(\exp i\alpha S_n^+)$, $v_n(\beta) = E(\exp i\beta S_n^-)$, Spitzer finds the relation (cf. [1, Eqn. (6.5)])

$$(1) \qquad \sum_{n=0}^{\infty} \phi_n(\alpha, \beta) t^n = \exp\left[\sum_{n=1}^{\infty} \frac{t^n}{n}(u_n(\alpha) + v_n(\beta) - 1)\right];$$

note that the distributions of $S_n^{\pm}$ are determined by that of $S_n$, so that by matching coefficients of $t^n$ we can read off the desired information. For $\beta = 0$ formula (1) specializes to

$$(1\text{a}) \qquad \sum_{n=0}^{\infty} \phi_n(\alpha) t^n = \exp \sum_{n=1}^{\infty} \frac{t^n}{n} u_n(\alpha),$$

$\phi_n(\alpha) = \phi_n(\alpha, 0) = E(\exp i\alpha R_n)$ (cf. [1, Eqn. (3.5)]).[1] Spitzer's derivation is based on a remarkable combinatorial lemma, Theorem 2.2 of [1]. The purpose of this note is to reverse the procedure: in §2 we give a direct derivation of (a variant of) (1), and in §3 we sketch the proof that the combinatorial lemma follows from (1a). The essential tool is the equation

$$\exp \log (e - g) = e - g,$$

suitably interpreted.

2. **The generating function.** Let $\zeta_n(\rho, \sigma) = E(\exp i(\rho R_n + \sigma S_n))$ and $\psi_n(\rho, \sigma) = E(\exp i(\rho S_n^+ + \sigma S_n))$. In this section we are going to deduce the formula

Received by the editors September 23, 1957 and, in revised form, May 27, 1958.

[1] The referee points out that formulas analogous to (1a) were found by Pollaczek [3] for the special case when the $X$'s possess a moment generating function.

$$\sum_{n=0}^{\infty} t^n \zeta_n(\rho, \sigma) = \exp \sum_{n=1}^{\infty} \frac{t^n}{n} \psi_n(\rho, \sigma). \tag{2}$$

That this is equivalent to (1) is seen by making the identifications $\rho = \alpha + \beta$, $\sigma = -\beta$ so that $\zeta_n(\rho, \sigma) = \phi_n(\rho + \sigma, -\sigma) = \phi_n(\alpha, \beta)$, and observing through an easy calculation that $\psi_n(\rho, \sigma) = u_n(\rho + \sigma) + v_n(-\sigma) - 1 = u_n(\alpha) + v_n(\beta) - 1$.

In order to prove (2) we consider the two-dimensional process $(R_n, S_n)$, starting from $R_0 = S_0 = 0$. The process is clearly Markovian, since its evolution is determined from the relations[2]

$$S_{n+1} = S_n + X_{n+1},$$
$$R_{n+1} = \max (R_n, S_n + X_{n+1}).$$

Let $g_n(r, s)$ denote the joint distribution function of $(R_n, S_n)$. Then the definitions imply that

$$g_{n+1}(r, s) = \int g_n(r, r \wedge s - x) d\, Pr\, (X < x) \tag{3}$$

in which $Pr(X < x)$ denotes the common distribution of the $X_n$ and $r \wedge s = \min (r, s)$; we shall also presently write $r \vee s$ for $\max (r, s)$.

In order to analyze the relation (3) we introduce $\mathfrak{M}$, the Banach algebra of linear combinations of distribution functions (i.e. the algebra of bounded Borel signed measures) over the plane, with convolution as multiplication and norm given by total variation. Let $e$ denote the identity of $\mathfrak{M}$, namely unit mass concentrated at the origin, and let $f$ be the common distribution function of the pairs $(0, X_n)$. For $g \in \mathfrak{M}$ we define $\exp g$ and $\log (e - g)$ by the usual Maclaurin series, assuming $\|g\| < 1$ for the latter; clearly $\exp \log (e - g) = e - g$. (The powers appearing in all series are of course repeated convolutions.) Next define linear operations $F$ and $P$ on $\mathfrak{M}$ by $Fg = fg$ and $(Pg)(r, s) = g(r, r \wedge s)$, $g \in \mathfrak{M}$; $P$ has the effect of projecting all mass to the left of the diagonal $D$: $s = r$ horizontally onto $D$, and leaving mass to the right of $D$ alone. For bounded Borel functions $h$ we note the relation

$$\iint h(r, s) d(Pg)(r, s) = \iint h(r \vee s, s) dg(r, s). \tag{4}$$

---

[2] The Markovian character of $(R_n, S_n)$ makes its study easier than that of the seemingly simpler but definitely non-Markovian process $R_n$; see however Spitzer [2], especially Eqn. (2.5), where it is remarked that each $R_n$ has the same distribution as $M_n$ defined recursively by $M_{n+1} = (M_n + X_{n+1})^+$, a Markov process.

Both $P$ and $F$ are continuous, in fact have norm one. $P$ is a projection leaving $e$ fixed; both $P\mathfrak{M}$ and $(I-P)\mathfrak{M}$ are closed subalgebras. Consequently for every $g$ we have $\exp Pg \in P\mathfrak{M}$ and $\exp (I-P)g = e+(I-P)g_1$ for some $g_1$.

Now we note that (3) is equivalent to $g_{n+1}=PFg_n$, and hence $g_n = (PF)^n e$, since $g_0 = e$. Therefore for $|t|<1$ we have

$$\sum_{n=0}^{\infty} t^n g_n = (I - tPF)^{-1}e. \tag{5}$$

We claim now that the right side of (5) equals $\exp\{-P \log (e-tf)\}$, from which it follows that

$$\sum_{n=0}^{\infty} t^n g_n = \exp \sum_{n=1}^{\infty} \frac{t^n}{n} P(f^n). \tag{6}$$

In order to establish the claim put $g=(I-tPF)^{-1}e$, so that $g - tPFg = e$. Then also $Pg - tPFg = e$ and therefore

$$Pg = g, \tag{7}$$

$$P[(e - tf)g] = e. \tag{8}$$

Conversely, if any $g^*$ satisfies (7) and (8) then already $g^* = g$. Now it is easy to see that this is the case for $g^* = \exp\{-P \log (e-tf)\}$; (7) is immediate because $g^* \in \exp P\mathfrak{M} \subseteq P\mathfrak{M}$, while (8) is established by the equations

$$\begin{aligned}(e - tf)g^* &= \exp \{\log (e - tf) - P \log (e - tf)\} \\ &= \exp \{(I - P) \log (e - tf)\} \\ &= e + (I - P)g_1\end{aligned}$$

for some $g_1$; applying $P$ then yields (8), and hence (6) is established.

It remains to prove (2). To do this take the Fourier-Stieltjes transform, $\mathfrak{T}$, of both sides of (6). For the left side we get

$$\begin{aligned}\mathfrak{T}(\sum t^n g_n) &= \sum t^n \mathfrak{T}(g_n) \\ &= \sum t^n \zeta_n(\rho, \sigma),\end{aligned}$$

and for the right side, since $\mathfrak{T}$ is not only continuous and linear but also multiplicative on $\mathfrak{M}$,

$$\mathfrak{T}\left(\exp \sum \frac{t^n}{n} P(f^n)\right) = \exp \sum \frac{t^n}{n} \mathfrak{T}P(f^n).$$

Then using (4) we have

$$\mathfrak{T}P(f^n) = \int\int \exp i(\rho r + \sigma s) d(Pf^n)(r, s)$$
$$= \int\int \exp i(\rho(r \vee s) + \sigma s) df^n(r, s)$$
$$= \int \exp i(\rho(0 \vee s) + \sigma s) dPr(S_n < s)$$
$$= \psi_n(\rho, \sigma)$$

completing the proof of (2).

3. **The combinatorial lemma.** In conclusion we sketch a proof that Spitzer's combinatorial lemma (Theorem 2.2 of [1]) follows from the formula (1a). The effect of this observation, combined with the direct proof of (1) and hence of (1a), is to show that simple Banach-algebraic formulas yield rather deep-seeming combinatorial facts, under suitable specialization.

Let $x=(x_1, x_2, \cdots, x_n)$ be a fixed $n$-tuple of real numbers. In the notation of [1] we wish to show that the two sets of numbers $[S(\sigma x)]$ and $[T(\tau x)]$ are identical, even counting multiplicities. (See [1] for the definitions of $S$ and $T$.)

In order to deduce this from (1a) let $p_1, p_2, \cdots, p_n$ be an arbitrary set of "probabilities," i.e. nonnegative real numbers adding up to one. Let $X$ be a random variable taking values $x_1, x_2, \cdots, x_n$ with the above probabilities, and write down the formula (1a) for the sequence of partial sums of independent copies of $X$. From both sides of the resulting expression extract the coefficients of $t^n$; they are equal, and are easily seen to be polynomials in the $p_i$, homogeneous of degree $n$. Because of the homogeneity they must be identically equal, that is, the conditions on the signs and on the sum of the $p_i$ can be disregarded. Therefore the coefficients of the product term $p_1p_2 \cdots p_n$ on the two sides of the polynomial equation must be equal. These coefficients are readily seen to be $n!\sum_\sigma \exp i\alpha S(\sigma x)$ and $n!\sum_\tau \exp i\alpha T(\tau x)$ respectively. The combinatorial lemma follows at once.

## References

**1.** F. Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. vol. 82 (1956) pp. 323–339.

**2.** ———, *The Wiener-Hopf equation whose kernel is a probability density*, Duke Math. J. vol. 24 (1947) pp. 327–343.

**3.** F. Pollaczek, *Fonctions caracteristiques de certaines répartitions définies au moyen de la notion d'ordre. Application à la théorie des attentes*, C. R. Acad. Sci. Paris vol. 234 (1952) pp. 2334–2336.

University of Michigan