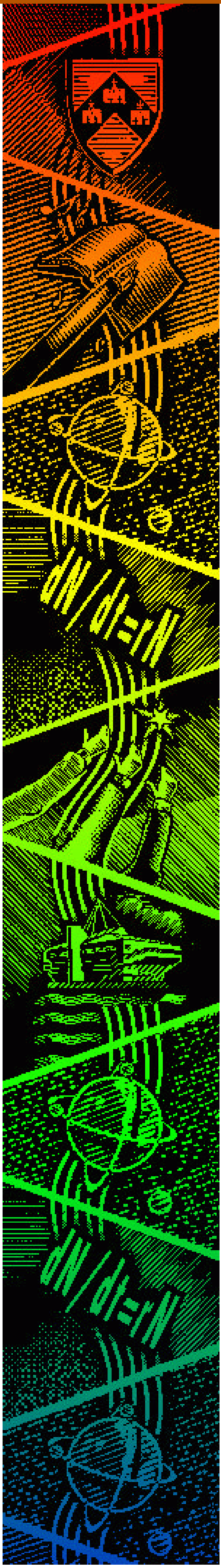

THE UNIVERSITY of York

Discussion Papers in Economics

No. 2000/06

The Density of the Sufficient Statistics for a Gaussian AR(1) Model in terms of Generalized Functions

by

Giovanni Forchini

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

The Density of the Sufficient Statistics for a Gaussian AR(1) Model in Terms of Generalized Functions

G. Forchini*
University of York

First version: November 1999
This version: January 2000

Abstract

This paper derives the exact joint distribution of the minimal sufficient statistics in the first-order AR(1) model with Gaussian errors and zero start-up value. The results are fundamental to an exact distribution theory for the statistics that are typically of interest in this model.

Key words and phrases: Gaussian AR(1) model, Generalized functions,

**Address for Correspondence:* Department of Economics and Related Studies, University of York, Heslington, York YO1 5DD, UK; e-mail: gf7@york.ac.uk

1 Introduction

The simple Gaussian AR(1) model

$$\begin{aligned} y_t &= \rho y_{t-1} + \varepsilon_t \\ t &= 1, 2, ..., T \\ \varepsilon_t &\sim NID\left(0, \sigma^2\right), y_0 = 0 \end{aligned} \qquad (1)$$

has attracted an enormous interest in both the statistical and econometric literature over many years. Despite the simplicity of the model, however, and despite many efforts, very little progress has been made on an exact (fixed T) distribution theory for the statistics that are typically of interest in (1) (estimators of ρ and σ^2, test statistics, etc.). Probably the best account of the results that have been won in this context is still the book by T. W. Anderson (1971).

Recently there has been a new interest in the exact distribution theory for the Gaussian AR(1) model. The first two moments of the maximum likelihood estimator for ρ are given by Sawa (1978), Nankervis and Savin (1988) and Vinod and Shenton (1996). Exact expressions for the statistical curvature, and the covariance matrix of the minimal sufficient statistics are derived by van Garderen (1997). Forchini (1998) has studied the properties of the exact density of the ordinary least squares estimator of ρ in (1).

The joint density of $y = (y_1, ..., y_T)'$ in (1) is

$$\text{pdf}\left(y; \rho, \sigma^2\right) = \left(2\pi\sigma^2\right)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[s_1 - 2\rho p + \rho^2 s_0\right]\right\}, \qquad (2)$$

where $s_0 = \sum_{t=1}^{T-1} y_t^2$, $s_1 = \sum_{t=1}^{T} y_t^2$, and $p = \sum_{t=1}^{T} y_t y_{t-1}$ are the minimal sufficient statistics for the model. Since the density of y is in the (curved) exponential family, the density of (s_0, s_1, p) belongs to the exponential family as well (Lehmann (1986), Section 2.7), and has precisely the form

$$\text{pdf}\left(s_0, s_1, p; \rho, \sigma^2\right) = \left(2\pi\sigma^2\right)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left[s_1 - 2\rho p + \rho^2 s_0\right]\right\} k\left(s_0, s_1, p\right),$$

where the function $k(s_0, s_1, p)$ does not depend on the parameters (ρ, σ^2). Note also that the joint density of (s_0, s_1, p) can be written as

$$\operatorname{pdf}\left(s_0, s_1, p|\rho, \sigma^2\right) = \left(\sigma^2\right)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma^2}\left(s_1 - 2\rho p + \rho^2 s_0\right)\right\} \exp\left\{\frac{s_1}{2}\right\}$$
$$\operatorname{pdf}\left(s_0, s_1, p|\rho = 0, \sigma^2 = 1\right)$$

so that attention can be focused on the joint density of (s_0, s_1, p) for $\rho = 0$ and $\sigma^2 = 1$. The functional form of $\operatorname{pdf}(s_0, s_1, p|\rho = 0, \sigma^2 = 1)$ is not known, yet, and an expression for it in terms of generalized functions (a summary of the theory of generalized functions with all the results of interest is in the Appendix) will be provided.

2 Joint density of the sufficient statistics

Our approach is based on the further decomposition

$$\operatorname{pdf}\left(s_0, s_1, p|\rho = 0, \sigma^2 = 1\right) = \operatorname{pdf}\left(p|s_0, s_1; \rho = 0, \sigma^2 = 1\right) \cdot \operatorname{pdf}\left(s_0, s_1|\rho = 0, \sigma^2 = 1\right). \tag{3}$$

The density of (s_0, s_1) is

$$\operatorname{pdf}\left(s_0, s_1|\rho = 0, \sigma^2 = 1\right) = \frac{\left(s_1 - s_0\right)^{-\frac{1}{2}} s_0^{\frac{T-1}{2}-1} \exp\left\{-\frac{s_1}{2}\right\} \exp\left\{-s_0\right\}}{2^{\frac{T}{2}} \pi^{\frac{1}{2}} \Gamma\left(\frac{T-1}{2}\right)}, \tag{4}$$

so that the conditional density of p given s_0 and s_1, with $s_0 s_1 \geq p^2$ for $\rho = 0$ and $\sigma^2 = 1$, must to be derived. This can be done by first deriving the characteristic function of p given (s_0, s_1), and then inverting it. This inversion constitutes the main problem because the characteristic function is written as an infinite series for which a closed form cannot be found, and which cannot be inverted term by term. However, if the characteristic function is interpreted as a *generalized function*, then term by term inversion is allowed.

2.1 The characteristic function of p given (s_0, s_1)

In matrix notation p can be written as

$$\begin{aligned} p &= \frac{1}{2} y' A_T y \\ &= \frac{1}{2} \bar{y}' A_{T-1} \bar{y} + y_T e'_{T-1} \bar{y} \end{aligned}$$

where $\bar{y} = (y_1, ..., y_{T-1})'$, and

$$A_T = \frac{1}{2} \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

is a $T \times T$ matrix and e_{T-1} is the $T-1 \times 1$ vector $e_{T-1} = (0, 0, ..., 0, 1)'$.

Setting $\bar{y} = v s_0^{\frac{1}{2}}$, $s_0 = \bar{y}'\bar{y} > 0$, $v = \bar{y} (\bar{y}'\bar{y})^{-\frac{1}{2}}$, $v \in S_{T-1}$, and $y_T = (s_1 - s_0)^{1/2} h$, $h \in O(1)$ delivers

$$p = \frac{1}{2} s_0 v' A_{T-1} v + h s_0^{\frac{1}{2}} (s_1 - s_0)^{\frac{1}{2}} e'_{T-1} v,$$

where $O(k)$ denote the group of $k \times k$ orthogonal matrices, and S_k is the set of k-dimensional vectors satisfying $v'v = 1$. Given $\rho = 0$ and $\sigma^2 = 1$, h and v are uniformly distributed on $O(k)$ and S_{T-1} respectively, and are mutually independent and are also independent of s_0 and s_1. Thus, the characteristic function of p given s_0 and s_1 is

$$\phi_{p|(s_0,s_1)}(t) = \int_{S_{T-1}} \int_{O(1)} \exp\left\{ -\frac{1}{2} it s_0 v' A_{T-1} v - ith s_0^{\frac{1}{2}} (s_1 - s_0)^{\frac{1}{2}} e'_{T-1} v \right\} (dv)(dh) \tag{5}$$

where (dv) and (dh) are the *normalized* Haar measures on S_{T-1} and $O(1)$ (Muirhead (1982)).

Integrating out h,

$$\int_{O(1)} \exp\left\{ -it s_0^{\frac{1}{2}} (s_1 - s_0)^{\frac{1}{2}} e'_{T-1} v h \right\} (dh) = \cos\left[t s_0^{\frac{1}{2}} (s_1 - s_0)^{\frac{1}{2}} e'_{T-1} v \right],$$

yields the conditional characteristic function of p,

$$\phi_{p|(s_0,s_1)}(t) = \int_{S_{T-1}} \exp\left\{-\frac{1}{2}its_0v'A_{T-1}v\right\} \cos\left[ts_0^{\frac{1}{2}}(s_1-s_0)^{\frac{1}{2}}e'_{T-1}v\right](dv). \qquad (7)$$

The integral over S_{T-1} is invariant under the transformation of A_{T-1} to $H'A_{T-1}H$, where $H = \text{diag}(1,h,1,h,....) \in O(T-1)$, $h \in O(1)$ (note that we could also choose $H = \text{diag}(h,1,h,1,....)$). This is the same as transforming $-\frac{1}{2}its_0v'A_{T-1}v$ to $-\frac{1}{2}its_0v'A_{T-1}vh$, $h \in O(1)$. Therefore, averaging the exponential over $O(1)$ as before, it can be shown that

$$\int_{O(1)} \exp\left\{-\frac{1}{2}its_0v'A_{T-1}v\right\}(dh) = \cos\left[ts_0v'A_{T-1}v\right],$$

therefore the conditional characteristic function (7) simplifies to

$$\phi_{p|(s_0,s_1)}(t) = \int_{S_{T-1}} \cos\left[ts_0v'A_{T-1}v\right] \cos\left[ts_0^{\frac{1}{2}}(s_1-s_0)^{\frac{1}{2}}e'_{T-1}v\right](dv). \qquad (8)$$

Expanding the cosines in power series and integrating term by term, the following uniformly convergent series can be obtained

$$\phi_{p|(s_0,s_1)}(t) = \sum_{n=0}^{\infty}\sum_{j=0}^{\infty} \frac{s_0^{2n+j}(s_1-s_0)^j c_{n.j}}{(2j)!(2n)!}(it)^{2n+2j}, \qquad (9)$$

where

$$c_{n.j} = \int_{S_{T-1}} (v'A_{T-1}v)^{2n}(e'_{T-1}v)^{2j}(v'dv). \qquad (10)$$

To evaluate (10), note that $v \in S_{T-1}$ can be written as $v = e'_{T-1}H$ where $H \in H(T-1)$ has the vector v as its last column. So the integral over S_{T-1} can be interpreted as an integral over $O(T-1)$, and

$$\begin{aligned} c_{n.j} &= \int_{O(T-1)} C_{[2n]}\left(e_{T-1}e'_{T-1}H'A_{T-1}H\right) \\ &\quad C_{[j]}\left(e_{T-1}e'_{T-1}H'e_{T-1}e'_{T-1}H\right)(dH) \\ &= \frac{C_{[2n+j]}^{[2n],[j]}\left(e_{T-1}e'_{T-1}, e_{T-1}e'_{T-1}\right) C_{[2n+j]}^{[2n],[j]}\left(A_{T-1}, e_{T-1}e'_{T-1}\right)}{C_{[2n+j]}\left(I_{T-1}\right)} \end{aligned}$$

where (dH) is the normalized invariant measure on the orthogonal group, $C_{[k]}(\cdot)$ denotes a top order invariant polynomial (Muirhead (1982)), $C_{[j+k]}^{[j],[k]}(.,.)$ is an invariant polynomial with two matrix arguments (Davis (1979), Chikuse and Davis (1986)), and $[k]$ indicates the top order partition of the integer k.

Since

$$C_{[2n+j]}^{[2n],[j]}\left(e_{T-1}e'_{T-1}, e_{T-1}e'_{T-1}\right) = 1$$

and

$$C_{[2n+j]}\left(I_{T-1}\right) = \frac{\left(\frac{T-1}{2}\right)_{2n+j}}{\left(\frac{1}{2}\right)_{2n+j}},$$

$c_{n,j}$ is

$$c_{n.j} = \frac{\left(\frac{1}{2}\right)_{2n+j}}{\left(\frac{T-1}{2}\right)_{2n+j}} C_{[2n+j]}^{[2n],[j]}\left(A_{T-1}, e_{T-1}e'_{T-1}\right). \tag{11}$$

Chikuse (1987) gives methods for evaluating the top-order Davis polynomials.

Summing by diagonals, $\phi_{p|(s_0,s_1)}(t)$ in (9) can be written as

$$\phi_{p|(s_0,s_1)}(t) = \sum_{n=0}^{\infty} a_n(s_0, s_1)(it)^{2n}, \tag{12}$$

where

$$a_n(s_0, s_1) = \sum_{j=0}^{n} \frac{\left(\frac{1}{2}\right)_{2n-j} s_0^{2n-j}(s_1 - s_0)^j}{\left(\frac{T-1}{2}\right)_{2n-j}(2j)!(2n-2j)!} C_{[2n-j]}^{[2n-2j],[j]}\left(A_{T-1}, e_{T-1}e'_{T-1}\right). \tag{13}$$

The conditional characteristic function is real and even, therefore the density of p given s_0 and s_1 must be symmetric.

2.2 Inversion

The conditional density of p given (s_0, s_1) is obtained by inverting the conditional characteristic function

$$\text{pdf}\left(p|s_0, s_1; \rho = 0, \sigma^2 = 1\right) = \frac{1}{2\pi}\int_{\mathbb{R}} \exp\{ipt\} \sum_{n=0}^{\infty} a_n(s_0, s_1)(it)^{2n}\, dt \tag{14}$$

$-\sqrt{s_0 s_1} \leq p \leq \sqrt{s_0 s_1}$. Term by term integration is not allowed, because the integral $\frac{1}{2\pi}\int_{\mathbb{R}} \exp\{ipt\}(it)^{2n}\, dt$ is not convergent. However, termwise inversion can be justified by using the theory of generalized functions (see the Appendix for a summary and Zemanian (1965) for more details). In other words the density of p given s_0 and s_1 is regarded as a functional, $\langle \mathrm{pdf}\,(p|s_0, s_1; 0, 1)\,, \varphi\rangle$ mapping from a space of "smooth" functions $\varphi\,(p)$ into the real numbers. These smooth functions will be used to "test" the quantity $\mathrm{pdf}\,(p|s_0, s_1; \rho = 0, \sigma^2 = 1)$ in the sense that this is evaluated by averaging it with respect to the testing functions in a finite interval about p,

$$p \mapsto \langle \mathrm{pdf}\,(p|s_0, s_1; 0, 1)\,, \varphi\rangle = \int_{[a,b]} \mathrm{pdf}\,(x|s_0, s_1; 0, 1)\,\varphi\,(x)\,dx,$$

$p \in [a, b]$. Using different testing functions φ, it can infer what the quantity $\mathrm{pdf}\,(p|s_0, s_1; 0, 1)$ looks like. The advantage of this approach is that what is "tested" does not need to be a function at all, and that many operations which are not allowed for functions are legitimate for generalized functions (Zemanian (1965)).

According to Theorem 1 in the Appendix, a series converging pointwise is convergent as a generalized function, and, thus, the inverse Fourier transform can be taken term by term, and the resulting series is convergent to a generalized function, too (Theorem 2 in the Appendix). Therefore,

$$\begin{aligned}\mathrm{pdf}\left(p|s_0, s_1; \rho = 0, \sigma^2 = 1\right) &= \sum_{n=0}^{\infty} a_n\,(s_0, s_1)\,\frac{1}{2\pi}\int_{\mathbb{R}} \exp\{ipt\}\,(it)^{2n}\,dt, \\ &= \sum_{n=0}^{\infty} a_n\,(s_0, s_1)\,\delta^{(2n)}\,(p)\,,\end{aligned}$$

where $\delta^{(m)}$ denotes the derivative of order m of the "delta function" (see the Appendix), $-\sqrt{s_1 s_0} \leq p \leq \sqrt{s_1 s_0}$, and $a_n\,(s_0, s_1)$ is defined in (13).

The joint density of (p, s_0, s_1) is, thus

$$\mathrm{pdf}\left(p, s_0, s_1|\rho = 0, \sigma^2 = 1\right) = \frac{\pi^{\frac{1}{2}}}{2^{\frac{T}{2}}\Gamma\left(\frac{T-1}{2}\right)} \tag{15}$$

$$(s_1 - s_0)^{-\frac{1}{2}}\, s_0^{\frac{T-1}{2}-1} \exp\left\{-\frac{s_1}{2}\right\} \exp\{-s_0\} \sum_{n=0}^{\infty} a_n\,(s_0, s_1)\,\delta^{(2n)}\,(p)$$

$$
\begin{aligned}
s_1 &> s_0 > 0, \\
-\sqrt{s_1 s_0} &\leq p \leq \sqrt{s_1 s_0}.
\end{aligned}
$$

It is important to emphasize again, that the expression (15) for the joint density of the sufficient statistics is not a function, but a generalized function.

Remarks.

1. Let $S(f) = \{t \in \mathbf{R} : f(t) \neq 0\}$ be the support of f. The support of $\delta^{(k)}(t)$ is $\{0\}$ for all $k \geq 0$. However, only a finite linear combination of the delta function and its derivatives has singular support (i.e. consisting of a point only). Therefore, expression (15) does not have a singular support.

2. The joint density of (p, s_0, s_1) depends on $|p|$ because $\delta^{(2j)}(-p) = \delta^{(2j)}(p)$, and is thus symmetric in p.

3. $\mathrm{pdf}(p, s_0, s_1)$ is an "ultradistribution", i.e. a continuous linear functional on the space of testing functions whose Fourier transforms are in the space of complex-valued functions that have continuous derivatives of all order and are zero outside some finite interval.

4. It is difficult to obtain the marginal densities of the statistics of interest from the joint density of the minimal sufficient statistics, because it is not clear whether integration and summation can be interchanged.

5. The technique used can be generalized to the case where $y_0 \neq 0$, and possibly to multivariate models.

6. Generalized functions might be used to simplify the derivation of asymptotic expansions. This point is currently under investigation.

7. The idea of deriving the density of p conditional on (s_0, s_1) can be used to find the density of other tests statistics. For example, the maximum likelihood estimator of ρ, can be written as

$$r = \frac{p}{s_0} = \frac{1}{2} v' A_{T-1} v + h \left(\frac{s_1}{s_0} - 1 \right)^{\frac{1}{2}} e'_{T-1} v$$

and the t-test statistic as $t = (r - \rho) \left(\frac{s_1}{s_0} - r^2 \right)^{-\frac{1}{2}} (T-2)^{\frac{1}{2}}$.

3 Concluding remarks

The result given in equation (3) provides the basis for an exact distribution theory for the maximum likelihood estimators for ρ and σ^2, as well as for other statistics that are of interest, such as test statistics for various hypotheses about ρ.

It seems plausible that the methods used in this paper might be capable of delivering analogous exact results for higher-order models, with formulae generalizing equation (3). However, since the number of sufficient statistics in a higher-order model is larger, the procedure might not be so straightforward in such cases.

Acknowledgements. Thanks are due to Francesco Bravo, Patrick Marsh and an unknown referee for comments which improved the presentation of the paper.

4 References

Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, John Wiley & Sons, New York.

Chikuse, Y. (1987), Methods for Constructing Top Order Invariant Polynomials, *Econometric Theory*, 3, 195-207.

Chikuse, Y. and A.W. Davis (1986), A Survey on the Invariant Polynomials with Matrix Arguments in Relation to Econometric Distribution Theory, *Econometric Theory*, 2, 232-248.

Davis, A.W. (1979), Invariant Polynomials with Two Matrix Arguments Extending the Zonal Polynomials: Applications in Econometrics, *Annals of the Institute of Statistical Mathematics*, 38, Part A, 109-122.

Forchini, G. (1998), The Distribution of the Serial Correlation Coefficient, *Mimeo*

Hurwicz, L. (1950), Least-Squares Bias in Time Series, in: T.C. Koopmans (ed.), *Statistical Inference in Dynamic Economic Models*, Cowles Commission for Research in Economics, Monograph No. 10, Chapman & Hall, London.

Lehmann, E.L. (1986), *Testing Statistical Hypotheses*, (2nd Ed.), John Wiley and Sons, New York.

Muirhead, R.J. (1982), *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc. New York.

Nankervis J.C. and N.W. Savin (1988), The Exact moments of the Least Squares Estimator for the Autoregressive Model: Corrections and Extensions, *Journal of Econometrics*, 37, 381-388.

Sawa, T. (1978), The Exact Moments of the Least Squares Estimator for the Autoregressive Model, *Journal of Econometrics*, 8, 159-172.

van Garderen, K.J. (1997), Exact Geometry of Explosive Autoregressive Models, CORE discussion paper No. 9768.

Vinod H.D. and L.R. Shenton (1996), Exact Moments for Autoregressive and Random Walk Models for a Zero or Stationary Initial Value, *Econometric Theory*, 12, 481-499.

Zemanian, A.H. (1965), *Distribution Theory and Transform Analysis*, Dover Publications Inc., New York.

5 Appendix: Review of Some Theory of Generalized Functions

There are two ways of evaluating a function. The first one is the standard one; $f(x)$ is evaluated at each particular point x: $x \to f(x)$. The second one requires an averaging of $f(x)$ with respect to another function $\phi(x)$, say: $x \to \int_a^b f(t)\,\phi(t)\,dt$, $x \in [a,b]$. The last approach leads to generalized functions.

Let $\mathcal{D}$ be the space of (*testing functions*) complex-valued functions $\phi(t)$ that have continuous derivatives of all order and are zero outside some finite interval. A *generalized function* $\langle f, \phi \rangle$ is a functional f on the space $\mathcal{D}$ that satisfies (i) (linearity) $\langle f, \alpha\phi_1 + \beta\phi_2 \rangle = \alpha \langle f, \phi_1 \rangle + \beta \langle f, \phi_2 \rangle$ for any ϕ_1 and ϕ_2 in $\mathcal{D}$ and any complex numbers α and β; and (ii) (continuity) for any sequence of testing functions $\{\phi_n\}_{n=1}^{\infty}$ that converges in $\mathcal{D}$ to $\phi(t)$ the sequence of numbers $\{\langle f, \phi_n \rangle\}_{n=1}^{\infty}$ converges to $\langle f, \phi \rangle$. The space of generalized functions defined on $\mathcal{D}$ is denoted by $\mathcal{D}'$.

Example 1. Let $f(t)$ be a locally integrable function (i.e. Lebesgue integrable over every finite set). A regular generalized function can be generated by $\langle f, \phi \rangle = \int_{\mathbf{R}} f(t)\,\phi(t)\,dt$. Note that if $\langle f, \phi \rangle = \langle g, \phi \rangle$ then $f = g$ a.e..

Example 2 The delta function, δ, is defined by the equation $\langle \delta, \phi \rangle = \int_{\mathbf{R}} \delta(t)\,\phi(t)\,dt = \phi(0)$. Note that the delta function is not a function. If $f(t)$ is a piecewise continuous function such that $\int_{\mathbf{R}} |f(t)|\,dt < \infty$ and $\int_{\mathbf{R}} f(t)\,dt = 1$ then $xf(xt) \to \delta(x)$ as $x \to \infty$.

The derivative f' of a distribution f is defined by $\langle f', \phi \rangle = -\langle f, \phi' \rangle$. In general the k-th derivative is defined by $\left\langle f^{(n)}, \phi \right\rangle = \left\langle f, (-1)^n \phi^{(n)} \right\rangle$.

Example 3. The first derivative of the delta function δ' is defined by $\langle \delta', \phi \rangle = -\langle \delta, \phi' \rangle = -\phi'(0)$.

The delta function can be approximated by the function $\gamma_a(t) = \zeta\left(\frac{t}{a}\right) / \int_{-a}^{a} \zeta\left(\frac{x}{a}\right) dx$, where $\zeta(t) = \begin{cases} 0 & \text{if } |t| \geq 1 \\ \exp\left\{\frac{1}{t^2-1}\right\} & \text{if } |t| < 1 \end{cases}$, as $a \to 0$. The derivatives of the delta function can be approximate by the derivative of $\gamma_a(t)$.

Definition 1 *The sequence of distributions* $\{f_n\}_{n=1}^{\infty}$ *converges in* $\mathcal{D}'$ *if, for every* ϕ *in* $\mathcal{D}$*, the sequence of numbers* $\{\langle f_n, \phi\rangle\}_{n=1}^{\infty}$ *converges. If* $\{f_n\}_{n=1}^{\infty}$ *converges in* $\mathcal{D}'$ *to the functional* f*, then* f *is also a distribution.*

Theorem 1 *(Zemanian (1965), Theorem 2.3-1).Let* $\{f_n(t)\}_{n=1}^{\infty}$ *be a sequence of locally integrable functions that converges pointwise almost everywhere to the function* $f(t)$ *and let all the functions* $f_n(t)$ *be bounded in magnitude by a locally integrable function. Then,* $f(t)$ *is locally integrable and the corresponding sequence of regular generalized functions* $\{f_n(t)\}_{n=1}^{\infty}$ *converges in* $\mathcal{D}'$ *to the regular generalized function* $f(t)$*.*

Note that if $\{f_n(t)\}_{n=1}^{\infty}$ is a sequence of locally integrable functions which converge uniformly on $\mathbb{R}$ then it converges as a generalized function.

Example 4. Let $f_n(t) = \sum_{i=0}^{n} \frac{(\alpha t)^i}{i!}$. Then $f_n(t) \to \sum_{i=0}^{\infty} \frac{(\alpha t)^i}{i!} = \exp\{\alpha t\}$.

Let $\mathcal{F}$ denote the Fourier transform. Then, the Fourier transform of a generalized function f is defined by $\langle \mathcal{F}(f), \phi\rangle = \langle f, \mathcal{F}(\phi)\rangle$. Let $\mathcal{Z}$ be the space of testing functions whose Fourier transforms are in $\mathcal{D}$, and denote $\mathcal{Z}'$ the space of continuous linear functionals on $\mathcal{Z}$ (*ultradistributions*). Then the following result holds.

Theorem 2 *(Zemanian (1965), Theorem 7.8-1). The Fourier transformation is a continuous linear mapping of* $\mathcal{D}'$ *onto* $\mathcal{Z}'$*. Hence if* $\sum_{v=1}^{\infty} g_v$ *converges in* $\mathcal{D}'$ *to* g*, then* $\mathcal{F}(g) = \sum_{v=1}^{\infty} \mathcal{F}(g_v)$*, where the last series converges in* $\mathcal{Z}'$*. The inverse Fourier transformation has the same properties as a mapping of* $\mathcal{Z}'$ *onto* $\mathcal{D}'$*.*