Fractional Time Evolution

R. Hilfer

ICA-1, Universität Stuttgart

70569 Stuttgart, Germany

Institut fÃijr Physik, Universität Mainz

55099 Mainz, Germany

from:
Applications of Fractional Calculus in Physics,
R. Hilfer (ed.), World Scientific, Singapore, 2000, page 89

ISBN 981-02-3457-0

Contents

CHAPTER 2

Fractional Time Evolution

[page 89, §0]

1. Introduction

[89.0.1] A large number of problems in theoretical physics, including Schrödingers, Maxwell and Newtons equations, can be formulated as initial value problems for dynamical evolution equations of the form

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = \mathrm{B} f(t) \tag{1}$$

where $t \in \mathbb{R}$ denotes time and B is an operator on a Banach space. *[89.0.2]* Depending on the initial data $f(0) = f_0$ describing the state or observable of the system at time $t = 0$ the problem is to find the state or observable $f(t)$ of the system at later times $t > 0$. [a]

[89.1.1] Many authors, mostly driven by the needs of applied problems, have considered generalizations of equation (1) of the form

$$\frac{\text{"}\mathrm{d}^\alpha\text{"}}{\mathrm{d}t^\alpha} f(t) = \mathrm{B} f(t) \tag{2}$$

in which the first order time derivative $\mathrm{d}/\mathrm{d}t$ is replaced with a certain fractional time derivative "$\mathrm{d}^\alpha/\mathrm{d}t^\alpha$" of order $\alpha > 0$ (see e.g. **[1]**–**[24]** and the Chapters IV–VIII in this book). *[89.1.2]* A number of fundamental questions are raised by such a replacement. *[89.1.3]* In order to appreciate these it is useful to recall that the appearance of $\mathrm{d}/\mathrm{d}t$ in eq. (1) reflects not only a basic symmetry of nature but also the basic principle of locality. *[89.1.4]* Of course, the symmetry in question is time translation invariance. *[89.1.5]* Remember that

$$\frac{\mathrm{d}}{\mathrm{d}s} f(s) = \lim_{t\to 0} \frac{f(s) - f(s-t)}{t} = -\lim_{t\to 0} \frac{\mathcal{T}(t) f(s) - f(s)}{t} \tag{3}$$

[a] In classical mechanics the states are points in phase space, the observables are functions on phase space, and the operator B is specified by a vector field and Poisson brackets. In quantum mechanics (with finitely many degrees of freedom) the states correspond to rays in a Hilbert space, the observables to operators on this space, and the operator B to the Hamiltonian. In field theories the states are normalized positive functionals on an algebra of operators or observables, and then B becomes a derivation on the algebra of observables. The equations (1) need not be first order in time. An example is the initial-value problem for the wave equation for $g(t, x)$

$$\frac{\partial^2 g}{\partial t^2} = c^2 \frac{\partial^2 g}{\partial x^2}$$

in one dimension. It can be recast into the form of eq. (1) by introducing a second variable h and defining

$$f = \begin{pmatrix} g \\ h \end{pmatrix} \quad \text{and} \quad \mathrm{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} c \frac{\partial}{\partial x}$$

[page 90, §0] identifies $-\mathrm{d}/\mathrm{d}t$ as the infinitesimal generator of time translations [b] defined by

$$\mathcal{T}(t)f(s) = f(s-t). \tag{4}$$

[90.0.1] Equation (2) abandons $\mathcal{T}(t)$ as the general time evolution, and this raises the question what replaces eq. (4), and how a fractional derivative can arise as the generator of a physical time evolution. *[90.0.2]* Most workers in fractional calculus have avoided these questions, and my purpose in this chapter is to review and discuss an answer provided recently in [**6, 7, 8, 9, 10, 11**].

[90.1.1] Derivatives of fractional order $0 < \alpha \leq 1$ were found to emerge quite generally as the infinitesimal generators of coarse grained macroscopic time evolutions given by [**6, 7, 8, 9, 10, 11**]

$$\mathrm{T}_\alpha(t)f(t_0) = \int_0^\infty \mathcal{T}(s)f(t_0)h_\alpha\left(\frac{s}{t}\right)\frac{\mathrm{d}s}{t} \tag{5}$$

where $t \geq 0$ and $0 < \alpha \leq 1$. *[90.1.2]* Explicit expressions for the kernels $h_\alpha(x)$ for all $0 < \alpha \leq 1$ are given in eq. (69) below. *[90.1.3]* It is the main objective of this chapter to show that (in a certain sense) all macroscopic time evolutions have the form of eq. (5), and that fractional time derivatives are their infinitesimal generators.

[90.2.1] Given the great difference between $\mathrm{T}_\alpha(t)$ in eq. (5) and $\mathrm{T}_1(t) = \mathcal{T}(t)$ in eq. (4) it becomes clear that basic issues, such as irreversibility, translation symmetry, or the meaning of stationarity are inevitably involved when proposing fractional dynamics. *[90.2.2]* Let me therefore advance the basic postulate that all time evolutions of physical systems are irreversible. *[90.2.3]* Obviously this *law of irreversibility* must be considered to be an empirical law of nature equal in rank to the law of energy conservation. *[90.2.4]* Reversible behaviour is an idealization. *[90.2.5]* Its validity or applicability in physical experiments depends on the degree to which the system can be isolated (or decoupled) from its past history and its environment. *[90.2.6]* According to this view the irreversible flow of time is more fundamental than the time reversal symmetry of Newtons or other equations. *[90.2.7]* My starting point is therefore that for a general time evolution operator $\mathrm{T}(t)$ the evolution parameter t is not a time instant (which could be positive or negative), but a duration, which cannot be negative.

[90.3.1] An immediate consequence of the postulated law of irreversibility is that the classical irreversibility problem of theoretical physics becomes reversed.

---

[b] A simple translation with unit "speed" reflects the idea of time "flowing" uniformly with constant velocity. This idea is embodied in measuring time by comparison with periodic processes (clocks). A competing idea, related to the flow of time represented by eq. (5), is to measure time by comparison with nonperiodic clocks such as decay or aging processes.

[page 91, §0] *[91.0.1]* Now the theoretical task is not to explain how irreversibility arises from reversible evolution equations, but how seemingly reversible equations arise as idealizations from an underlying irreversible time evolution. *[91.0.2]* A possible explanation is provided by the present theory based on eq. (5). *[91.0.3]* It turns out that the case $\alpha = 1$ in eq. (5) is of predominant mathematical and physical importance, because it is in a quantifiable sense a strong universal attractor. *[91.0.4]* In this case the kernel $h_1(x)$ becomes

$$h_1(x) = \lim_{\alpha \to 1^-} h_\alpha(x) = \delta(x-1), \tag{6}$$

and the time evolution $\mathrm{T}_1(t) = \mathcal{T}(t)$ in (5) reduces to a simple translation as in eq.(4). *[91.0.5]* $T_1(t)$ with $t \geq 0$ is a representation of the time semigroup $(\mathbb{R}_+, +)$. *[91.0.6]* It can be extended to one of the full group $(\mathbb{R}, +)$. *[91.0.7]* This is not possible for T_α with $0 < \alpha < 1$. *[91.0.8]* The physical interpretation of α is seen from $\operatorname{supp} h_\alpha = \mathbb{R}_+$ for $\alpha \neq 1$ and $\operatorname{supp} h_1 = \{1\}$ for $\alpha = 1$. *[91.0.9]* Hence the parameter α classifies and quantifies the influence of the past history. *[91.0.10]* Small values of α correspond to a strong influence of the past history. *[91.0.11]* For $\alpha = 1$ the influence of the past history is minimal in the sense that it enters only through the present state.

[91.1.1] The basic result in eq. (5) was given in [**6**] and subsequently rationalized within ergodic theory by investigating the recurrence properties of induced automorphisms on subsets of measure zero [**9, 10, 11**]. *[91.1.2]* In these investigations the existence of a recurrent subset of measure zero had to be assumed. *[91.1.3]* Such an assumption becomes plausible from observations in low dimensional chaotic systems (see e.g. [**25, 26**] and Chapter V). *[91.1.4]* A rigorous proof for any given dynamical system, however, appears difficult, and it is therefore of interest to rederive the emergence of $\mathrm{T}_\alpha(t)$ from a different, and more general, approach.

2. Foundations

2.1. Basic Desiderata for Time Evolutions

[91.2.1] The following basic requirements define a time evolution in this chapter.

(1) Semigroup
[91.2.2] A time evolution is a pair $(\{\mathrm{T}_\tau(t) : 0 \leq t < \infty\}, (B_\tau, \|\cdot\|))$ where $\mathrm{T}_\tau(t) = \mathrm{T}(t\tau)$ is a semigroup of operators $\{\mathrm{T}(t) : 0 \leq t < \infty\}$ mapping the Banach space $(B_\tau(\mathbb{R}), \|\cdot\|)$ of functions $f_\tau(s) = f(s\tau)$ on $\mathbb{R}$ to itself. *[91.2.3]* The argument $t \geq 0$ of $\mathrm{T}_\tau(t)$ represents a time duration, the argument $s \in \mathbb{R}$ of $f_\tau(s)$ a time instant. *[91.2.4]* The index $\tau > 0$ indicates the units (or scale) of time. *[91.2.5]* Below, τ will again be frequently suppressed to simplify the notation. *[91.2.6]* The elements $f_\tau(s) = f(s\tau) \in B_\tau$ represent observables or

[page 92, §0]
the state of a physical system as function of the time coordinate $s \in \mathbb{R}$. *[92.0.1]* The semigroup conditions require

$$\mathrm{T}_\tau(t_1)\mathrm{T}_\tau(t_2)f_\tau(t_0) = \mathrm{T}_\tau(t_1+t_2)f_\tau(t_0) \tag{7}$$

$$\mathrm{T}_\tau(0)f_\tau(t_0) = f_\tau(t_0) \tag{8}$$

for $t_1, t_2 > 0$, $t_0 \in \mathbb{R}$ and $f_\tau \in B_\tau$. *[92.0.2]* The first condition may be viewed as representing the unlimited divisibility of time.

(2) Continuity
[92.0.3] The time evolution is assumed to be strongly continuous in t by demanding

$$\lim_{t\to 0} \|\mathrm{T}(t)f - f\| = 0 \tag{9}$$

for all $f \in B$.

(3) Homogeneity
[92.0.4] The homogeneity of the time coordinate requires commutativity with translations

$$\mathcal{T}(t_1)\mathrm{T}(t_2)f(t_0) = \mathrm{T}(t_2)\mathcal{T}(t_1)f(t_0) \tag{10}$$

for all $t_2 > 0$ and $t_0, t_1 \in \mathbb{R}$. *[92.0.5]* This postulate allows to shift the origin of time and it reflects the basic symmetry of time translation invariance.

(4) Causality
[92.0.6] The time evolution operator should be causal in the sense that the function $g(t_0) = (\mathrm{T}(t)f)(t_0)$ should depend only on values of $f(s)$ for $s < t_0$.

(5) Coarse Graining
[92.0.7] A time evolution operator $\mathrm{T}(t)$ should be obtainable from a coarse graining procedure. *[92.0.8]* A precise definition of coarse graining is given in Definition 2.3 below. *[92.0.9]* The idea is to combine a time average $\frac{1}{t}\int_{s-t}^{s} f(t')\,\mathrm{d}t'$ in the limit $t, s \to \infty$ with a rescaling of s and t.

[92.0.10] While the first four requirements are conventional the fifth requires comment. *[92.0.11]* Averages over long intervals may themselves be timedependent on much longer time scales. *[92.0.12]* An example would be the position of an atom in a glass. *[92.0.13]* On short time scales the position fluctuates rapidly around a well defined average position. *[92.0.14]* On long time scales the structural relaxation processes in the glass can change this average position. *[92.0.15]* The purpose of any coarse graining procedure is to connect microscopic to macroscopic scales. *[92.0.16]* Of course, what is microscopic

[page 93, §0] depends on the physical situation. *[93.0.1]* Any microscopic time evolution may itself be viewed as macroscopic from the perspective of an underlying more microscopic theory. *[93.0.2]* Therefore it seems physically necessary and natural to demand that a time evolution should generally be obtainable from a coarse graining procedure.

2.2. Evolutions, Convolutions and Averages

[93.1.1] There is a close connection and mathematical similarity between the simplest time evolution $\mathrm{T}(t) = \mathcal{T}(t)$ and the operator $\mathrm{M}(t)$ of time averaging defined as the mathematical mean

$$\mathrm{M}(t)f(s) = \frac{1}{t}\int_{s-t}^{s} f(y)\,\mathrm{d}y, \tag{11}$$

where $t > 0$ is the length of the averaging interval. *[93.1.2]* Rewriting this formally as

$$\mathrm{M}(t)f(s) = \frac{1}{t}\int_{0}^{t} f(s-y)\,\mathrm{d}y = \frac{1}{t}\int_{0}^{t} \mathcal{T}(y)f(s)\,\mathrm{d}y \tag{12}$$

exhibits the relation between $\mathrm{M}(t)$ and $\mathcal{T}(t)$. *[93.1.3]* It shows also that $\mathrm{M}(t)$ commutes with translations (see eq. (10)).

[93.2.1] A second even more suggestive relationship between $\mathrm{M}(t)$ and $\mathcal{T}(t)$ arises because both operators can be written as convolutions. *[93.2.2]* The operator $\mathrm{M}(t)$ may be written as

$$\begin{aligned}\mathrm{M}(t)f(s) &= \frac{1}{t}\int_{0}^{t} f(s-y)\,\mathrm{d}y = \int_{-\infty}^{\infty} f(s-y)\frac{1}{t}\chi_{[0,1]}\left(\frac{y}{t}\right)\,\mathrm{d}y \\ &= \int_{0}^{s} f(s-y)\frac{1}{t}\chi_{[0,1]}\left(\frac{y}{t}\right)\,\mathrm{d}y,\end{aligned} \tag{13}$$

where the kernel

$$\chi_{[0,1]}(x) = \begin{cases} 1 & \text{for } x \in [0,1] \\ 0 & \text{for } x \notin [0,1] \end{cases} \tag{14}$$

is the characteristic function of the unit interval. *[93.2.3]* The Laplace convolution in the last line requires $t < s$. *[93.2.4]* The translations $\mathcal{T}(t)$ on the other hand may be *[page 94, §0]* written as

$$\mathcal{T}(t)f(s) = f(s-t) = \int_{-\infty}^{\infty} f(s-y)\frac{1}{t}\delta\left(\frac{y}{t}-1\right)\,\mathrm{d}y$$
$$= \int_{0}^{s} f(s-y)\frac{1}{t}\delta\left(\frac{y}{t}-1\right)\,\mathrm{d}y \tag{15}$$

where again $0 < t < s$ is required for the Laplace convolution in the last equation. *[94.0.1]* The similarity between eqs. (15) and (13) suggests to view the time translations $\mathcal{T}(t)$ as a degenerate form of averaging f over a single point. *[94.0.2]* The operators $\mathrm{M}(t)$ and $\mathcal{T}(t)$ are both convolution operators. *[94.0.3]* By Lebesgues theorem $\lim_{t\to 0}\mathrm{M}(t)f(s) = f(s)$ so that $\mathrm{M}(0)f(t) = f(t)$ in analogy with eq. (8) which holds for $\mathcal{T}(t)$. *[94.0.4]* However, while the translations $\mathcal{T}(t)$ fulfill eq. (7) and form a convolution semigroup whose kernel is the Dirac measure at 1, the averaging operators $\mathrm{M}(t)$ do not form a semigroup as will be seen below.

[94.1.1] The appearance of convolutions and convolution semigroups is not accidental. *[94.1.2]* Convolution operators arise quite generally from the symmetry requirement of eq. (10) above. *[94.1.3]* Let $L^p(\mathbb{R}^n)$ denote the Lebesgue spaces of p-th power integrable functions, and let $\mathcal{S}$ denote the Schwartz space of test functions for tempered distributions [**27**]. *[94.1.4]* It is well established that all bounded linear operators on $L^p(\mathbb{R}^n)$ commuting with translations (i.e. fulfilling eq. (10)) are of convolution type [**27**].

Theorem 2.1 *[94.1.5] Suppose the operator* $\mathrm{B}: L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)$, $1 \le p, q, \le \infty$ *is linear, bounded and commutes with translations. [94.1.6] Then there exists a unique tempered distribution* g *such that* $\mathrm{B}h = g * h$ *for all* $h \in \mathcal{S}$.

[94.2.1] For $p = q = 1$ the tempered distributions in this theorem are finite Borel measures. *[94.2.2]* If the measure is bounded and positive this means that the operator B can be viewed as a weighted averaging operator. *[94.2.3]* In the following the case $n = 1$ will be of interest. *[94.2.4]* A positive bounded measure μ on $\mathbb{R}$ is uniquely determined by its distribution function $\widetilde{\mu}: \mathbb{R} \to [0, 1]$ defined by

$$\widetilde{\mu}(x) = \frac{\mu(]-\infty, x[)}{\mu(\mathbb{R})}. \tag{16}$$

[94.2.5] The tilde will again be omitted to simplify the notation. *[94.2.6]* Physically a weighted average $\mathrm{M}(t;\mu)f(s)$ represents the measurement of a signal $f(s)$ using an apparatus with response characterized by μ and resolution $t > 0$. *[94.2.7]* Note that the resolution (length of averaging interval) is a duration and cannot be negative.

[page 95, §1]

Definition 2.1 (Averaging) *[95.1.1] Let μ be a (probability) distribution function on $\mathbb{R}$, and $t > 0$. [95.1.2] The weighted (time) average of a function f on $\mathbb{R}$ is defined as the convolution*

$$\mathrm{M}(t;\mu)f(s) = (f * \mu(\cdot/t))(s) = \int_{-\infty}^{\infty} f(s-s')\,\mathrm{d}\mu(s'/t) = \int_{-\infty}^{\infty} \mathcal{T}(s')f(s)\,\mathrm{d}\mu(s'/t) \quad (17)$$

whenever it exists. [95.1.3] The average is called causal if the support of μ is in $\mathbb{R}_+$. [95.1.4] It is called degenerate if the support of μ consists of a single point.

[95.2.1] The weight function or kernel $m(x)$ corresponding to a distribution $\mu(x)$ is defined as $m(x) = \mathrm{d}\mu/\mathrm{d}x$ whenever it exists.

[95.3.1] The averaging operator $\mathrm{M}(t)$ in eq. (11) corresponds to a measure with distribution function

$$\mu_\chi(x) = \begin{cases} 0 & \text{for } x \le 0 \\ x & \text{for } 0 \le x \le 1 \\ 1 & \text{for } x \ge 1 \end{cases} \quad (18)$$

while the time translation $\mathcal{T}(t)$ corresponds to the (Dirac) measure $\delta(x-1)$ concentrated at 1 with distribution function

$$\mu_\delta(x) = \begin{cases} 0 & \text{for } x < 1 \\ 1 & \text{for } x \ge 1. \end{cases} \quad (19)$$

[95.3.2] Both averages are causal, and the latter is degenerate.

[95.4.1] Repeated averaging leads to convolutions. *[95.4.2]* The convolution κ of two distributions μ, ν on $\mathbb{R}$ is defined through

$$\kappa(x) = (\mu * \nu)(x) = \int_{-\infty}^{\infty} \mu(x-y)\mathrm{d}\nu(y) = \int_{-\infty}^{\infty} \nu(x-y)\mathrm{d}\mu(y). \quad (20)$$

[95.4.3] The Fourier transform of a distribution is defined by

$$\mathcal{F}\{\mu(t)\}(\omega) = \widehat{\mu}(\omega) = \int_{-\infty}^{\infty} e^{i\omega t}\mathrm{d}\mu(t) = \int_{-\infty}^{\infty} e^{i\omega t} m(t)\,\mathrm{d}t \quad (21)$$

where the last equation holds when the distribution admits a weight function. *[95.4.4]* A sequence $\mu_n(x)$ of distributions is said to converge weakly to a limit $\mu(x)$, *[page 96, §0]* written as

$$\lim_{n\to\infty}\mu_n = \mu, \tag{22}$$

if

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f(x)\mathrm{d}\mu_n(x) = \int_{-\infty}^{\infty} f(x)\mathrm{d}\mu_(x) \tag{23}$$

holds for all bounded continuous functions f.

[96.1.1] The operators $\mathrm{M}(t)$ and $\mathcal{T}(t)$ above have positive kernels, and preserve positivity in the sense that $f \geq 0$ implies $\mathrm{M}(t)f \geq 0$. *[96.1.2]* For such operators one has

Theorem 2.2 *[96.1.3] Let* T *be a bounded operator on* $L^p(\mathbb{R})$, $1 \leq p < \infty$ *that is translation invariant in the sense that*

$$\mathrm{T}\mathcal{T}(t)f = \mathcal{T}(t)\mathrm{T}f \tag{24}$$

for all $t \in \mathbb{R}$ *and* $f \in L^p(\mathbb{R})$, *and such that* $f \in L^p(\mathbb{R})$ *and* $0 \leq f \leq 1$ *almost everywhere implies* $0 \leq \mathrm{T}f \leq 1$ *almost everywhere. [96.1.4] Then there exists a uniquely determined bounded measure* μ *on* $\mathbb{R}$ *with mass* $\mu(\mathbb{R}) \leq 1$ *such that*

$$\mathrm{T}f(t) = (\mu * f)(t) = \int_{-\infty}^{\infty} f(t-s)\mathrm{d}\mu(s) \tag{25}$$

PROOF. *[96.1.5]* For the proof see [**28**]. □

[96.1.6] The preceding theorem suggests to represent those time evolutions that fulfill the requirements 1.– 4. of the last section in terms of convolution semigroups of measures.

Definition 2.2 (Convolution semigroup) *[96.1.7] A family* $\{\mu_t : t > 0\}$ *of positive bounded measures on* $\mathbb{R}$ *with the properties that*

$$\mu_t(\mathbb{R}) \leq 1 \quad \text{for } t > 0, \tag{26}$$

$$\mu_{t+s} = \mu_t * \mu_s \quad \text{for } t, s > 0, \tag{27}$$

$$\delta = \lim_{t\to 0}\mu_t \tag{28}$$

is called a convolution semigroup of measures on $\mathbb{R}$.

[page 97, §1]
[97.1.1] Here δ is the Dirac measure at 0 and the limit is the weak limit. *[97.1.2]* The desired characterization of time evolutions now becomes

Corollary 2.1 *[97.1.3] Let* $\mathrm{T}(t)$ *be a strongly continuous time evolution fulfilling the conditions of homogeneity and causality, and being such that* $f \in L^p(\mathbb{R})$ *and* $0 \le f \le 1$ *almost everywhere implies* $0 \le \mathrm{T}f \le 1$ *almost everywhere. [97.1.4] Then* $\mathrm{T}(t)$ *corresponds uniquely to a convolution semigroup of measures* μ_t *through*

$$\mathrm{T}(t)f(s) = (\mu_t * f)(s) = \int_{-\infty}^{\infty} f(s-s')\mathrm{d}\mu_t(s') \tag{29}$$

with $\operatorname{supp}\mu_t \subset \mathbb{R}_+$ *for all* $t \ge 0$.

Proof. *[97.1.5]* Follows from Theorem 2.2 and the observation that $\operatorname{supp}\mu_t \cap \mathbb{R}_- \neq \emptyset$ would violate the causality condition. □

[97.2.1] Equation (29) establishes the basic convolution structure of the assertion in eq. (5). *[97.2.2]* It remains to investigate the requirement that $\mathrm{T}(t)$ should arise from a coarse graining procedure, and to establish the nature of the kernel in eq. (5).

2.3. Time Averaging and Coarse Graining

[97.3.1] The purpose of this section is to motivate the definition of coarse graining. *[97.3.2]* A first possible candidate for a coarse grained macroscopic time evolution could be obtained by simply rescaling the time in a microscopic time evolution as

$$\mathrm{T}_\infty(\bar{t})f(s) = \lim_{\tau\to\infty} \mathrm{T}_\tau(\bar{t})f(s) = \lim_{\tau\to\infty} \mathrm{T}(\tau\bar{t})f(s) = \lim_{\tau\to\infty} f(s-\tau\bar{t}) \tag{30}$$

where $0 < \bar{t} < \infty$ would be macroscopic times. *[97.3.3]* However, apart from special cases, the limit will in general not exist. *[97.3.4]* Consider for example a sinusoidal $f(t)$ oscillating around a constant. *[97.3.5]* Also, the infinite translation T_∞ is not an average, and this conflicts with the requirement above, that coarse graining should be a smoothing operation.

[97.4.1] A second highly popular candidate for coarse graining is therefore the averaging operator $\mathrm{M}(t)$. *[97.4.2]* If the limit $t \to \infty$ exists and $f(t)$ is integrable in the finite interval $[s_1, s_2]$ then the average

$$\overline{f} = \lim_{t\to\infty} \mathrm{M}(t)f(s_1) = \lim_{t\to\infty} \mathrm{M}(t)f(s_2) \tag{31}$$

is a number independent of the instant s_i. *[97.4.3]* Thus, if one wants to study the macroscopic time dependence of $\overline{f}$, it is necessary to consider a scaling limit in *[page 98, §0]* which also $s \to \infty$. *[98.0.1]* If the scaling limit $s, t \to \infty$ is performed such that $s/t = \overline{s}$ is constant, then

$$\lim_{\substack{t,s\to\infty \\ s=t\overline{s}}} \mathrm{M}(t)f(s) = \int_{\overline{s}-1}^{\overline{s}} f_\infty(z)\,\mathrm{d}z = \mathrm{M}(1)f_\infty(\overline{s}) \tag{32}$$

becomes again an averaging operator over the infinitely rescaled observable. *[98.0.2]* Now $\mathrm{M}(1)$ still does not qualify as a coarse grained time evolution because $\mathrm{M}(1)\mathrm{M}(1) \neq \mathrm{M}(2)$ as will be shown next.

[98.1.1] Consider again the operator $\mathrm{M}(t)$ defined in eq. (11). *[98.1.2]* It follows that

$$\mathrm{M}^2(t)f(s) = \left(\frac{1}{t}\chi_{[0,1]}\left(\frac{\cdot}{t}\right) * \frac{1}{t}\chi_{[0,1]}\left(\frac{\cdot}{t}\right) * f\right)(s) \tag{33}$$

and

$$\frac{1}{t^2}\int_0^x \chi_{[0,1]}\left(\frac{x-y}{t}\right)\chi_{[0,1]}\left(\frac{y}{t}\right)\,\mathrm{d}y = \begin{cases} 0 & \text{for } x \le 0 \\ \dfrac{x}{t^2} & \text{for } 0 \le x \le t \\ \dfrac{2}{t} - \dfrac{x}{t^2} & \text{for } t \le x \le 2t \\ 0 & \text{for } x \ge 2t. \end{cases} \tag{34}$$

[98.1.3] Thus twofold averaging may be written as

$$\mathrm{M}^2(t)f(s) = \int_0^s f(s-y)\frac{1}{t}\chi^{(2)}\left(\frac{y}{t}\right)\,\mathrm{d}y \tag{35}$$

where

$$\chi^{(2)}(x) = \begin{cases} x & \text{for } 0 \le x \le 1 \\ 2-x & \text{for } 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases} \tag{36}$$

is the new kernel. *[98.1.4]* It follows that $\mathrm{M}^2(t) \neq \mathrm{M}(2t)$, and hence the averaging operators $\mathrm{M}(t)$ do not form a semigroup.

[98.2.1] Although $\mathrm{M}^2(t) \neq \mathrm{M}(2t)$ the iterated average is again a convolution operator with support $[0, 2t]$ compared to $[0, t]$ for $\mathrm{M}(t)$. *[98.2.2]* Similarly, $\mathrm{M}^3(t)$ has support $[0, 3t]$. *[98.2.3]* This suggests to investigate the iterated average $\mathrm{M}^n(t)f(s)$ in a scaling limit $n, s \to \infty$. *[98.2.4]* The limit $n \to \infty$ smoothes the function by enlarging the

[page 99, §0] averaging window to $[0, nt]$, and the limit $s \to \infty$ shifts the origin to infinity. *[99.0.1]* The result may be viewed as a coarse grained time evolution in the sense of a time evolution on time scales "longer than infinitely long". *[99.0.2]* [c] It is therefore necessary to rescale s. *[99.0.3]* If the rescaling factor is called $\sigma_n > 0$ one is interested in the limit $n, s \to \infty$ with $\overline{s} = s/\sigma_n$ fixed, and $\sigma_n \to \infty$ with $n \to \infty$ and fixed $t > 0$

$$\lim_{\substack{n,s\to\infty\\ s=\sigma_n\overline{s}}} (\mathrm{M}(t)^n f)(s) = \lim_{n\to\infty} (\mathrm{M}(t)^n f)(\sigma_n \overline{s}) \tag{37}$$

whenever this limit exists. *[99.0.4]* Here $\overline{s} > 1$ denotes the macroscopic time.

[99.1.1] To evaluate the limit note first that eq. (11) implies

$$\mathrm{M}(t) f(\sigma_n \overline{s}) = \int_0^{\overline{s}} f_{\sigma_n}(\overline{s} - z) \frac{\sigma_n}{t} \chi_{[0,1]}\left(\frac{\sigma_n z}{t}\right) \mathrm{d}z \tag{38}$$

where $f_\tau(t) = f(t\tau)$ denotes the rescaled observable with a rescaling factor τ. *[99.1.2]* The n-th iterated average may now be calculated by Laplace transformation with respect to $\overline{s}$. *[99.1.3]* Note that

$$\mathcal{L}\left\{\frac{1}{c}\chi_{[0,1]}\left(\frac{x}{c}\right)\right\}(u) = \frac{1 - e^{-cu}}{cu} = E_{1,2}(-cu) \tag{39}$$

for all $c \in \mathbb{R}$, where $E_{1,2}(x)$ is the generalized Mittag-Leffler function defined as

$$E_{a,b}(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(ak + b)} \tag{40}$$

for all $a > 0$ and $b \in \mathbb{C}$. *[99.1.4]* Using the general relation

$$E_{a,b}(x) = \frac{1}{\Gamma(b)} + x E_{a,a+b}(x) \tag{41}$$

gives with eqs. (37) and (38)

$$\mathcal{L}\{\mathrm{M}(t)^n f(\sigma_n \overline{s})\}(\overline{u}) = \left(1 - \frac{t\overline{u}}{\sigma_n} E_{1,3}\left(-\frac{t\overline{u}}{\sigma_n}\right)\right)^n \frac{1}{\sigma_n} \mathcal{L}\{f(s)\}\left(\frac{\overline{u}}{\sigma_n}\right) \tag{42}$$

where $f(\overline{u})$ is the Laplace transform of $f(\overline{s})$. *[99.1.5]* Noting that $E_{1,3}(0) = 1/2$ it becomes apparent that a limit $n \to \infty$ will exist if the rescaling factors are

[c] The scaling limit was called "ultralong time limit" in [**10**]

[page 100, §0] chosen as $\sigma_n \sim n$. *[100.0.1]* With the choice $\sigma_n = \sigma n/2$ and $\sigma > 0$ one finds for the first factor

$$\lim_{n\to\infty}\left(1-\frac{2t\overline{u}}{n\sigma}E_{1,3}\left(-\frac{2t\overline{u}}{n\sigma}\right)\right)^n = e^{-t\overline{u}/\sigma}. \tag{43}$$

[100.0.2] Concerning the second factor assume that for each $\overline{u}$ the limit

$$\lim_{n\to\infty}\frac{2}{n}\mathcal{L}\left\{f(s)\right\}\left(\frac{2\overline{u}}{n}\right) = \overline{f}(\overline{u}) \tag{44}$$

exists and defines a function $\overline{f}(\overline{u})$. *[100.0.3]* Then

$$\lim_{n\to\infty}\frac{1}{\sigma_n}\mathcal{L}\left\{f(\overline{s})\right\}\left(\frac{\overline{u}}{\sigma_n}\right) = \frac{1}{\sigma}\overline{f}\left(\frac{\overline{u}}{\sigma}\right), \tag{45}$$

and it follows that

$$\lim_{n\to\infty}\mathcal{L}\left\{\mathrm{M}(t)^n f(\sigma_n\overline{s})\right\}(\overline{u}) = e^{-t\overline{u}/\sigma}\frac{1}{\sigma}\overline{f}\left(\frac{\overline{u}}{\sigma}\right). \tag{46}$$

[100.0.4] With $\bar{t} = t/\sigma$ Laplace inversion yields

$$\lim_{\substack{n,s\to\infty\\ s=\sigma_n\overline{s}}}(\mathrm{M}(t)^n f)(s) = \int_0^{\overline{s}}\overline{f}(\sigma\overline{s}-\sigma\overline{y})\delta(\overline{y}-\bar{t})\,\mathrm{d}\overline{y} = \overline{f}_\sigma(\overline{s}-\bar{t}). \tag{47}$$

[100.0.5] Using eq. (12) the result (47) may be expressed symbolically as

$$\lim_{\substack{n,s\to\infty\\ s/n=\sigma\overline{s}/2}}\left(\frac{1}{t}\int_0^t\mathcal{T}(y)\,\mathrm{d}y\right)^n f(s) = \overline{f}_\sigma(\overline{s}-\bar{t}) = \overline{\mathcal{T}}(\bar{t})\,\overline{f}_\sigma(\overline{s}) \tag{48}$$

with $\bar{t} = t/\sigma$. *[100.0.6]* This expresses the macroscopic or coarse grained time evolution $\overline{\mathcal{T}}(\bar{t})$ as the scaling limit of a microscopic time evolution $\mathcal{T}(t)$. *[100.0.7]* Note that there is some freedom in the choice of the rescaling factors σ_n expressed by the prefactor σ. *[100.0.8]* This freedom reflects the freedom to choose the time units for the coarse grained time evolution.

[100.1.1] The coarse grained time evolution $\overline{\mathcal{T}}(\bar{t})$ is again a translation. *[100.1.2]* The coarse grained observable $\overline{f}(\overline{s})$ corresponds to a microscopic average by virtue of the following result **[29]**.

[page 101, §1]

Proposition 2.1 *[101.1.1] If $f(x)$ is bounded from below and one of the limits*

$$\lim_{y\to\infty}\frac{1}{y}\int_0^y f(x)\,\mathrm{d}x$$

or

$$\lim_{z\to 0} z\int_0^\infty f(x)e^{-zx}\,\mathrm{d}x$$

exists then the other limit exists and

$$\lim_{y\to\infty}\frac{1}{y}\int_0^y f(x)\,\mathrm{d}x = \lim_{z\to 0} z\mathcal{L}\{f(x)\}(z). \tag{49}$$

[101.1.2] Comparison of the last relation with eq. (44) shows that $\overline{f}(\overline{s})$ is a microscopic average of $f(s)$. *[101.1.3]* While s is a microscopic time coordinate, the time coordinate $\overline{s}$ of $\overline{f}$ is macroscopic.

[101.2.1] The preceding considerations justify to view the time evolution $\overline{\mathcal{T}}(\bar{t})$ as a coarse grained time evolution. *[101.2.2]* Every observation or measurement of a physical quantity $f(s)$ requires a minimum duration t determined by the temporal resolution of the measurement apparatus. *[101.2.3]* The value $f(s)$ at the time instant s is always an average over this minimum time interval. *[101.2.4]* The averaging operator $\mathrm{M}(t)$ with kernel $\chi_{[0,1]}$ defined in equation (11) represents an idealized averaging apparatus that can be switched on and off instantaneously, and does not otherwise influence the measurement. *[101.2.5]* In practice one is usually confronted with finite startup and shutdown times and a nonideal response of the apparatus. *[101.2.6]* These imperfections are taken into account by using a weighted average with a weight function or kernel that differs from $\chi_{[0,1]}$. *[101.2.7]* The weight function reflects conditions of the measurement, as well as properties of the apparatus and its interaction with the system. *[101.2.8]* It is therefore of interest to consider causal averaging operators $\mathrm{M}(t;\mu)$ defined in eq. (17) with general weight functions. *[101.2.9]* A general coarse graining procedure is then obtained from iterating these weighted averages.

Definition 2.3 (Coarse Graining) *[101.2.10] Let μ be a probability distribution on $\mathbb{R}$, and $\sigma_n > 0$, $n \in \mathbb{N}$ a sequence of rescaling factors. A coarse graining limit is defined as*

$$\lim_{\substack{n,s\to\infty\\ s=\sigma_n\overline{s}}}(\mathrm{M}(t;\mu)^n f)(s) \tag{50}$$

[page 102, §0] whenever the limit exists. [102.0.1] The coarse graining limit is called causal if $\mathrm{M}(t;\mu)$ *is causal, i.e. if* $\operatorname{supp}\mu \subset \mathbb{R}_+$.

2.4. Coarse Graining Limits and Stable Averages

[102.1.1] The purpose of this section is to investigate the coarse graining procedure introduced in Definition 2.3. *[102.1.2]* Because the coarse graining procedure is defined as a limit it is useful to recall the following well known result for limits of distribution functions [**30**]. *[102.1.3]* For the convenience of the reader its proof is reproduced in the appendix.

Proposition 2.2 *[102.1.4] Let* $\mu_n(s)$ *be a weakly convergent sequence of distribution functions. [102.1.5] If* $\lim_{n\to\infty}\mu_n(s)=\mu(s)$, *where* $\mu(s)$ *is nondegenerate then for any choice of* $a_n>0$ *and* b_n *there exist* $a>0$ *and* b *such that*

$$\lim_{n\to\infty}\mu_n(a_nx+b_n)=\mu(ax+b). \tag{51}$$

[102.2.1] The basic result for coarse graining limits can now be formulated.

Theorem 2.3 (Coarse Graining Limit) *[102.2.2] Let* $f(s)$ *be such that the limit* $\lim_{a\to 0} a\widehat{f}(a\omega)=\widehat{\overline{f}}(\omega)$ *defines the Fourier transform of a function* $\overline{f}(s)$. *[102.2.3] Then the coarse graining limit exists and defines a convolution operator*

$$\lim_{\substack{n,s\to\infty\\ s=\sigma_n\overline{s}}}(\mathrm{M}(t;\mu)^n f)(s)=\int_{-\infty}^{\infty}\overline{f}(\overline{s}-\overline{s}')\,\mathrm{d}\nu(\overline{s}'/t;\mu) \tag{52}$$

if and only if for any $a_1,a_2>0$ *there are constants* $a>0$ *and* b *such that the distribution function* $\nu(x)=\nu(x;\mu)$ *obeys the relation*

$$\nu(a_1x)*\nu(a_2x)=\nu(ax+b). \tag{53}$$

PROOF. *[102.2.4]* In the previous section the coarse graining limit was evaluated for the distribution μ_χ from eq. (18) and the corresponding ν was found in eq. (47) to be degenerate. *[102.2.5]* A degenerate distribution ν trivially obeys eq. (53). *[102.2.6]* Assume therefore from now on that neither μ nor ν are degenerate.

[102.3.1] Employing equation (17) in the form

$$\mathrm{M}(t;\mu)f(\sigma_n\overline{s})=\int_{-\infty}^{\infty}f(\sigma_n\overline{s}-\sigma_n y)\mathrm{d}\mu(\sigma_n y/t) \tag{54}$$

[page 103, §0] one computes the Fourier transformation of $\mathrm{M}(t;\mu)^n f$ with respect to $\overline{s}$

$$\mathcal{F}\left\{\mathrm{M}(t;\mu)^n f(\sigma_n\overline{s})\right\}(\overline{\omega}) = \left[\widehat{\mu}\left(\frac{t\overline{\omega}}{\sigma_n}\right)\right]^n \frac{1}{\sigma_n}\widehat{f}\left(\frac{\overline{\omega}}{\sigma_n}\right). \tag{55}$$

[103.0.1] By assumption $\widehat{f}(\overline{\omega}/\sigma_n)/\sigma_n$ has a limit whenever $\sigma_n \to \infty$ with $n \to \infty$. *[103.0.2]* Thus the coarse graining limit exists and is a convolution operator whenever $[\widehat{\mu}(t\overline{\omega}/\sigma_n)]^n$ converges to $\widehat{\nu}(\overline{\omega})$ as $n \to \infty$. *[103.0.3]* Following [**30**] it will be shown that this is true if and only if the characterization (53) and $\sigma_n \to \infty$ with $n \to \infty$ apply. *[103.0.4]* To see that

$$\lim_{n\to\infty} \sigma_n = \infty \tag{56}$$

holds, assume the contrary. Then there is a subsequence σ_{n_k} converging to a finite limit. *[103.0.5]* Thus

$$|\widehat{\mu}(t\omega/\sigma_{n_k})|^{n_k} = |\widehat{\nu}(\omega)|(1+o(1)) \tag{57}$$

so that

$$|\widehat{\mu}(\omega)| = |\widehat{\nu}(\omega\sigma_{n_k}/t)|^{1/n_k}(1+o(1)) \tag{58}$$

for all ω. *[103.0.6]* As $n_k \to \infty$ this leads to $|\widehat{\mu}(\omega)| = 1$ for all ω and hence μ must be degenerate contrary to assumption.

[103.1.1] Next, it will be shown that

$$\lim_{n\to\infty} \frac{\sigma_{n+1}}{\sigma_n} = 1. \tag{59}$$

[103.1.2] From eq. (56) it follows that $\lim_{n\to\infty} |\widehat{\mu}(\omega/\sigma_n)| = 1$ and therefore

$$|\widehat{\mu}(t\omega/\sigma_n)|^n = |\widehat{\nu}(\omega)|(1+o(1)) \tag{60}$$

and

$$|\widehat{\mu}(t\omega/\sigma_{n+1})|^{n+1} = |\widehat{\nu}(\omega)|(1+o(1)). \tag{61}$$

Substituting ω by $\sigma_n\omega/\sigma_{n+1}$ in eq. (60) and by $\sigma_{n+1}\omega/\sigma_n$ in eq. (61) shows that

$$\lim_{n\to\infty}\left|\frac{\widehat{\nu}(\sigma_{n+1}\omega/\sigma_n)}{\widehat{\nu}(\omega)}\right| = \lim_{n\to\infty}\left|\frac{\widehat{\nu}(\sigma_n\omega/\sigma_{n+1})}{\widehat{\nu}(\omega)}\right| = 1. \tag{62}$$

[page 104, §0] *[104.0.1]* If $\lim_{n\to\infty}\sigma_{n+1}/\sigma_n \neq 1$ then there exists a subsequence of either (σ_{n+1}/σ_n) or (σ_n/σ_{n+1}) converging to a constant $A < 1$. *[104.0.2]* Therefore eq. (62) implies $\widehat{\nu}(\omega) = \widehat{\nu}(A\omega)$ which upon iteration yields

$$|\widehat{\nu}(\omega)| = |\widehat{\nu}(A^n\omega)|. \tag{63}$$

[104.0.3] Taking the limit $n \to \infty$ then gives $|\widehat{\nu}(0)| = 1$ implying that ν is degenerate contrary to assumption.

[104.1.1] Now let $0 < a_1 < a_2$ be two constants. *[104.1.2]* Because of (56) and (59) it is possible to choose for each $\varepsilon > 0$ and sufficiently large $n > n_0(\varepsilon)$ an index $m(n)$ such that

$$0 \leq \frac{\sigma_m}{\sigma_n} - \frac{a_2}{a_1} < \varepsilon. \tag{64}$$

[104.1.3] Consider the identity

$$\left[\widehat{\mu}\left(\frac{a_1 t\overline{\omega}}{\sigma_n}\right)\right]^{n+m} = \left[\widehat{\mu}\left(\frac{a_1 t\overline{\omega}}{\sigma_n}\right)\right]^{n}\left[\widehat{\mu}\left(\frac{\sigma_m}{\sigma_n}\frac{a_1 t\overline{\omega}}{\sigma_m}\right)\right]^{m}. \tag{65}$$

By hypothesis the distribution functions corresponding to $[\widehat{\mu}\,(t\overline{\omega}/\sigma_n)]^n$ converge to $\nu(\overline{s})$ as $n \to \infty$. *[104.1.4]* Hence each factor on the right hand side converges and their product converges to $\nu(a_1\overline{s}) * \nu(a_2\overline{s})$. *[104.1.5]* It follows that the distribution function on the left hand side must also converge. *[104.1.6]* By Proposition 2.2 there must exist $a > 0$ and b such that the left hand side differs from $\nu(\overline{s})$ only as $\nu(a\overline{s} + b)$.

[104.2.1] Finally the converse direction that the coarse graining limit exists for $\mu = \nu$ is seen to follow from eq. (53). *[104.2.2]* This concludes the proof of the theorem. □

[104.3.1] The theorem shows that the coarse graining limit, if it exists, is again a macroscopic weighted average $\mathrm{M}(t;\nu)$. *[104.3.2]* The condition (53) says that this macroscopic average has a kernel that is stable under convolutions, and this motivates the

Definition 2.4 (Stable Averages) *[104.3.3] A weighted averaging operator* $\mathrm{M}(t;\mu)$ *is called stable if for any* $a_1, a_2 > 0$ *there are constants* $a > 0$ *and* $b \in \mathbb{R}$ *such that*

$$\mu(a_1x) * \mu(a_2x) = \mu(ax + b) \tag{66}$$

holds.

[104.4.1] This nomenclature emphasizes the close relation with the limit theorems of probability theory [**30, 31**]. *[104.4.2]* The next theorem provides the explicit form for distribution functions satisfying eq. (66). *[104.4.3]* The proof uses Bernsteins theorem and hence requires the concept of complete monotonicity.

[page 105, §1]

Definition 2.5 *[105.1.1] A C^∞-function $f :]0,\infty[\to \mathbb{R}$ is called completely monotone if*

$$(-1)^n \frac{\mathrm{d}^n f}{\mathrm{d}x^n} \geq 0 \tag{67}$$

for all integers $n \geq 0$.

[105.2.1] Bernsteins theorem [**31**, p. 439] states that a function is completely monotone if and only if it is the the Laplace transform ($u > 0$)

$$\mu(u) = \mathcal{L}\left\{\mu(x)\right\}(u) = \int_0^\infty e^{-ux}\mathrm{d}\mu(x) = \int_0^\infty e^{-ux} m(x)\,\mathrm{d}x \tag{68}$$

of a distribution μ or of a density $m = \mathrm{d}\mu/\mathrm{d}x$.

[105.3.1] In the next theorem the explicit form of stable averaging kernels is found to be a special case of the general H-function. *[105.3.2]* Because the H-function will reappear in other results its general definition and properties are presented separately in Section 4.

Theorem 2.4 *[105.3.3] A causal average is stable if and only if its weight function is of the form*

$$h_\alpha(x; b, c) = \frac{1}{b^{1/\alpha}} h_\alpha\left(\frac{x-c}{b^{1/\alpha}}\right) = \frac{1}{\alpha(x-c)} H^{10}_{11}\left(\frac{b^{1/\alpha}}{x-c}\,\middle|\,\begin{matrix}(0,1)\\(0,1/\alpha)\end{matrix}\right) \tag{69}$$

where $0 < \alpha \leq 1$, $b > 0$ and $c \in \mathbb{R}$ are constants and $h_\alpha(x) = h_\alpha(s; 1, 0)$.

Proof. *[105.3.4]* Let $c = 0$ without loss of generality. *[105.3.5]* The condition (66) together with $\operatorname{supp}\mu \subset [0,\infty[$ defines one sided stable distribution functions [**31**]. *[105.3.6]* To derive the form (69) it suffices to consider condition (66) with $b = 0$. *[105.3.7]* Assume thence that for any $a_1, a_2 > 0$ there exists $a > 0$ such that

$$\mu(a_1 x) * \mu(a_2 x) = \mu(ax) \tag{70}$$

where the convolution is now a Laplace convolution because of the condition supp $\subset$ $[0,\infty[$. *[105.3.8]* Laplace tranformation yields

$$\mu(u/a_1)\mu(u/a_2) = \mu(u/a). \tag{71}$$

[105.3.9] Iterating this equation (with $a_1 = a_2 = 1$) shows that there is an n-dependent constant $a(n)$ such that

$$\mu(u)^n = \mu(u/a(n)) \tag{72}$$

[page 106, §0] and hence

$$\mu\left(\frac{u}{a(nm)}\right) = \mu(u)^{nm} = \mu\left(\frac{u}{a(n)}\right)^m = \mu\left(\frac{u}{(a(n)a(m)}\right). \tag{73}$$

[106.0.1] Thus $a(n)$ satisfies the functional equation

$$a(nm) = a(n)a(m) \tag{74}$$

whose solution is $a(n) = n^{1/\gamma}$ with some real constant written as $1/\gamma$ with hindsight. *[106.0.2]* Inserting $a(n)$ into eq.(72) and substituting the function $g(x) = \log\mu(x)$ gives

$$ng(u) = g(un^{-1/\gamma}). \tag{75}$$

[106.0.3] Taking logarithms and substituting $f(x) = \log g(e^x)$ this becomes

$$\log n + f(\log u) = f\left(\log u - \frac{\log n}{\gamma}\right). \tag{76}$$

[106.0.4] The solution to this functional equation is $f(x) = -\gamma x$. *[106.0.5]* Substituting back one finds $g(x) = x^{-\gamma}$ and therefore $\mu(u)$ is of the general form $\mu(u) = \exp(u^{-\gamma})$ with $\gamma \in \mathbb{R}$. *[106.0.6]* Now μ is also a distribution function. Its normalization requires $\mu(u = 0) = 1$ and this restricts γ to $\gamma < 0$. *[106.0.7]* Moreover, by Bernsteins theorem $\mu(u)$ must be completely monotone. *[106.0.8]* A completely monotone function is positive, decreasing and convex. *[106.0.9]* Therefore the power in the exponent must have a negative prefactor, and the exponent is restricted to the range $-1 \leq \gamma < 0$. *[106.0.10]* Summarizing, the Laplace transform $\mu(u)$ of a distribution satisfying (70) is of the form

$$\mu(u) = h_\alpha(u; b, 0) = e^{-bu^\alpha} \tag{77}$$

with $0 < \alpha \leq 1$ and $b > 0$. *[106.0.11]* Checking that $h_\alpha(u; b, 0)$ does indeed satisfy eq. (70) yields $a^{-\alpha} = a_1^{-\alpha} + a_2^{-\alpha}$ as the relation between the constants. *[106.0.12]* For the proof of the general case of eq. (66) see Refs. **[30, 31]**.

[106.1.1] To invert the Laplace transform it is convenient to use the relation

$$\mathcal{M}\left\{m(x)\right\}(s) = \frac{\mathcal{M}\left\{\mathcal{L}\left\{m(x)\right\}(u)\right\}(1-s)}{\Gamma(1-s)} \tag{78}$$

between the Laplace transform and the Mellin transform

$$\mathcal{M}\left\{m(x)\right\}(s) = \int_0^\infty x^{s-1} m(t)\,\mathrm{d}x \tag{79}$$

[page 107, §0] of a function $m(x)$. *[107.0.1]* Using the Mellin transform **[32]**

$$\mathcal{M}\left\{e^{-bx^{\alpha}}\right\}(s) = \frac{\Gamma(s/\alpha)}{\alpha b^{s/\alpha}} \tag{80}$$

valid for $\alpha > 0$ and $\operatorname{Re} s > 0$ it follows that

$$\mathcal{M}\left\{h_\alpha(x;b,0)\right\}(s) = \frac{1}{\alpha b^{(1-s)/\alpha}} \frac{\Gamma((1-s)/\alpha)}{\Gamma(1-s)}. \tag{81}$$

[107.0.2] The general relation $\mathcal{M}\left\{x^{-1}f(x^{-1})\right\}(s) = \mathcal{M}\left\{f(x)\right\}(1-s)$ then implies

$$\mathcal{M}\left\{x^{-1}h_\alpha(x^{-1};b,0)\right\}(s) = \frac{1}{\alpha b^{s/\alpha}} \frac{\Gamma(s/\alpha)}{\Gamma(s)} \tag{82}$$

which leads to

$$x^{-1}h_\alpha(x^{-1};b,0) = \frac{1}{\alpha} H^{10}_{11}\left(b^{1/\alpha}x \,\middle|\, \begin{matrix} (0,1) \\ (0,1/\alpha) \end{matrix} \right) \tag{83}$$

by identification with eq. (153) below. *[107.0.3]* Restoring a shift $c \neq 0$ yields the result of eq. (69). □

[107.0.4] Note that $h_\alpha(x) = h_\alpha(s;1,0)$ is the standardized form used in eq. (5). *[107.0.5]* It remains to investigate the sequence of rescaling factors σ_n. *[107.0.6]* For these one finds

Corollary 2.2 *[107.0.7] If the coarse graining limit exists and is nondegenerate then the sequence σ_n of rescaling factors has the form*

$$\sigma_n = n^{1/\alpha}\Lambda(n) \tag{84}$$

where $0 < \alpha \leq 1$ and $\Lambda(n)$ is slowly varying, i.e. $\lim_{n\to\infty} \Lambda(bn)/\Lambda(n) = 1$ for all $b > 0$ (see Chapter IX, Section 2.3).

Proof. **[33]** *[107.0.8]* Let $\widehat{\mu}_n(\omega) = \widehat{\mu}(\omega)^n$. *[107.0.9]* Then for all ω and any fixed k

$$|\widehat{\mu}_n(\omega/\sigma_n)| = e^{-b|\omega|^\alpha}(1+o(1)) = |\widehat{\mu}_{kn}(\omega/\sigma_{kn})|. \tag{85}$$

[107.0.10] On the other hand

$$|\widehat{\mu}_{kn}(\omega/\sigma_{kn})| = |\widehat{\mu}_n((\omega\sigma_n/\sigma_{kn})/\sigma_n)|^k = e^{-b|\omega|^\alpha}(1+o(1)) \tag{86}$$

where the remainder tends uniformly to zero on every finite interval. *[107.0.11]* Suppose that the sequence σ_n/σ_{kn} is unbounded so that there is a subsequence with $\sigma_{kn_j}/\sigma_{n_j} \to 0$. *[107.0.12]* Setting $\omega = \sigma_{kn_j}/\sigma_{n_j}$ in eq. (86) and using eq. (85) gives *[page 108, §0]* $\exp(-bk) = 1$ which cannot be satisfied because $b, k > 0$. *[108.0.1]* Hence σ_n/σ_{kn} is bounded. *[108.0.2]* Now the limit $n \to \infty$ in eqs. (85) and (86) gives

$$e^{-b|\omega|^\alpha} = e^{-bk|\omega|^\alpha(\sigma_n/\sigma_{kn})^\alpha}(1 + o(1)). \tag{87}$$

[108.0.3] This requires that

$$\lim_{n\to\infty} \frac{\sigma_{kn}}{\sigma_n} = k^{1/\alpha} \tag{88}$$

implying eq. (84) by virtue of the Characterization Theorem 2.2 in Chapter IX. *[108.0.4]* (For more information on slow and regular variation see Chapter IX and references therein). □

2.5. Macroscopic Time Evolutions

[108.1.1] The preceding results show that a coarse graining limit is characterized by the quantities (α, b, c, Λ). *[108.1.2]* These quantities are determined by the coarsening weight μ. *[108.1.3]* The following result, whose proof can be found in [**33**, p. 85], gives their relation with the coarsening weight.

Theorem 2.5 (Universality Classes of Time Evolutions) *[108.1.4] In order that a causal coarse graining limit based on* $\mathrm{M}(t;\mu)$ *gives rise to a macroscopic average with* $h_\alpha(x; b, c)$ *it is necessary and sufficient that* $\widehat{\mu}(\omega)$ *behaves as*

$$\log \widehat{\mu}(\omega) = ic\omega - b|\omega|^\alpha \Lambda(\omega) \tag{89}$$

in a neighbourhood of $\omega = 0$*, and that* $\Lambda(\omega)$ *is slowly varying for* $\omega \to 0$. *[108.1.5] In case* $0 < \alpha \le 1$ *the rescaling factors can be chosen as*

$$\sigma_n^{-1} = \inf\{\omega > 0 : |\omega|^\alpha \Lambda(\omega) = b/n\} \tag{90}$$

while the case $\alpha > 1$ *reduces to the degenerate case* $\alpha = 1$.

[108.2.1] The preceding theorem characterizes the domain of attraction of a universality class of time evolutions. *[108.2.2]* Summarizing the results gives a characterization of macroscopic time evolutions arising from coarse graining limits.

Theorem 2.6 (Macroscopic Time Evolution) *[108.2.3] Let* $f(s)$ *be such that the limit* $\lim_{a\to 0} a\widehat{f}(a\omega) = \widehat{\overline{f}}(\omega)$ *defines the Fourier transform of a function* $\overline{f}(s)$. *[108.2.4] If* $\mathrm{M}(t;\mu)$ *is a causal average whose coarse graining limit exists with* α, b, c *as* *[page 109, §0] in the preceding theorem then*

$$\lim_{\substack{n,s\to\infty\\ s=\sigma_n\overline{s}}}(\mathrm{M}(t;\mu)^n f)(s) = \int\limits_{\overline{c}}^{\infty} \overline{f}(\overline{s}-y)h_\alpha\left(\frac{y}{\overline{t}}\right)\frac{\mathrm{d}y}{\overline{t}} = \int\limits_{\overline{c}}^{\infty} \overline{\mathcal{T}}_y\overline{f}(\overline{s})h_\alpha\left(\frac{y}{\overline{t}}\right)\frac{\mathrm{d}y}{\overline{t}}$$
$$= \mathrm{M}(\overline{t};h_\alpha)\overline{f}(\overline{s}-\overline{c}) = \overline{\mathrm{T}}_\alpha(\overline{t})\overline{f}(\overline{s}-\overline{c}) \tag{91}$$

defines a family of one parameter semigroups $\overline{\mathrm{T}}_\alpha(\overline{t})$ with parameter $\overline{t} = t^\alpha b$ indexed by α. [109.0.1] Here $\overline{\mathcal{T}}_{\overline{t}}\overline{f}(\overline{s}) = \overline{f}(\overline{s}-\overline{t})$ denotes the translation semigroup, and $\overline{c} = c/(tb)^{1/\alpha}$ is a constant.

Proof. *[109.0.2]* Noting that $\operatorname{supp} h_\alpha(x) \subset \mathbb{R}_+$ and combining Theorems 2.3 and 2.4 gives

$$\lim_{\substack{n,s\to\infty\\ s=\sigma_n\overline{s}}}(\mathrm{M}(t;\mu)^n f)(s) = \int_c^\infty \overline{f}(\overline{s}-\overline{s}')\frac{1}{tb^{1/\alpha}}h_\alpha\left(\frac{\overline{s}'-c}{tb^{1/\alpha}}\right)\mathrm{d}\overline{s}' = \overline{\mathrm{T}}_\alpha(\overline{t})\overline{f}(\overline{s}-\overline{c}) \tag{92}$$

where $0 < \alpha \le 1$, $b > 0$ and $c \in \mathbb{R}$ are the constants from theorem 2.4 and the last equality defines the operators $\overline{\mathrm{T}}_\alpha(\overline{t})$ with $\overline{t} = t^\alpha b$ and $\overline{c} = c/(tb)^{1/\alpha}$. *[109.0.3]* Fourier transformation then yields

$$\mathcal{F}\left\{(\overline{\mathrm{T}}_\alpha(\overline{t})\overline{f})(\overline{s}-\overline{c})\right\}(\overline{\omega}) = e^{-ic\overline{\omega}-\overline{t}(i\overline{\omega})^\alpha}, \tag{93}$$

and the semigroup property (7) follows from

$$\mathcal{F}\left\{(\overline{\mathrm{T}}_\alpha(\overline{t}_1)\overline{\mathrm{T}}_\alpha(\overline{t}_2)\overline{f})(\overline{s}-\overline{c})\right\}(\overline{\omega}) = e^{-ic\overline{\omega}-\overline{t}_1(i\overline{\omega})^\alpha-\overline{t}_2(i\overline{\omega})^\alpha}$$
$$= \mathcal{F}\left\{(\overline{\mathrm{T}}_\alpha(\overline{t}_1+\overline{t}_2)\overline{f})(\overline{s}-\overline{c})\right\}(\overline{\omega}) \tag{94}$$

by Fourier inversion. *[109.0.4]* Condition (8) is checked similarly. □

[109.0.5] The family of semigroups $\overline{\mathrm{T}}_\alpha(\overline{t})$ indexed by α that can arise from coarse graining limits are called *macroscopic time evolutions*. *[109.0.6]* These semigroups are also holomorphic, strongly continuous and equibounded (see Chapter III).

[109.1.1] From a physical point of view this result emphasizes the different role played by $\overline{s}$ and $\overline{t}$. *[109.1.2]* While $\overline{s}$ is the macroscopic time coordinate whose values are $\overline{s} \in \mathbb{R}$, the duration $\overline{t} > 0$ is positive. *[109.1.3]* If the dimension of a microscopic time duration t is [s], then the dimension of the macroscopic time duration $\overline{t}$ is [s$^\alpha$].

[page 110, §0]

2.6. Infinitesimal Generators

[110.0.1] The importance of the semigroups $\overline{\mathrm{T}}_\alpha(\overline{t})$ for theoretical physics as universal attractors of coarse grained macroscopic time evolutions seems not to have been noticed thus far. *[110.0.2]* This is the more surprising as their mathematical importance for harmonic analysis and probability theory has long been recognized [**31, 34, 35, 28**]. *[110.0.3]* The infinitesimal generators are known to be fractional derivatives [**31, 35, 36, 37**]. *[110.0.4]* The infinitesimal generators are defined as

$$\mathrm{A}_\alpha \overline{f}(\overline{s}) = \lim_{\overline{t}\to 0} \frac{\overline{\mathrm{T}}_\alpha(\overline{t})\overline{f}(\overline{s}) - \overline{f}(\overline{s})}{\overline{t}}. \tag{95}$$

[110.0.5] For more details on semigroups and their infinitesimal generators see Chapter III.

[110.1.1] Formally one calculates A_α by applying direct and inverse Laplace transformation with $\overline{c} = 0$ in eq. (91) and using eq. (77)

$$\begin{aligned}
\mathrm{A}_\alpha \overline{f}(\overline{s}) &= \lim_{\overline{t}\to 0} \frac{1}{2\pi i} \int_{\eta - i\infty}^{\eta + i\infty} e^{\overline{s}\overline{u}} \left(\frac{e^{-\overline{t}\overline{u}^\alpha} - 1}{\overline{t}} \right) \overline{f}(\overline{u})\,\mathrm{d}\overline{u} \\
&= \frac{1}{2\pi i} \int_{\eta - i\infty}^{\eta + i\infty} e^{\overline{s}\overline{u}} \lim_{\overline{t}\to 0} \left(\frac{e^{-\overline{t}\overline{u}^\alpha} - 1}{\overline{t}} \right) \overline{f}(\overline{u})\,\mathrm{d}\overline{u} \\
&= -\frac{1}{2\pi i} \int_{\eta - i\infty}^{\eta + i\infty} e^{\overline{s}\overline{u}} \overline{u}^\alpha \overline{f}(\overline{u})\,\mathrm{d}\overline{u}.
\end{aligned} \tag{96}$$

[110.1.2] The result can indeed be made rigorous and one has

Theorem 2.7 *[110.1.3] The infinitesimal generator A_α of the macroscopic time evolutions $\overline{\mathrm{T}}_\alpha(\overline{t})$ is related to the infinitesimal generator $\mathrm{A} = -\mathrm{d}/\mathrm{d}\overline{t}$ of $\overline{\mathcal{T}}_{\overline{t}}$ through*

$$\begin{aligned}
\mathrm{A}_\alpha \overline{f}(\overline{s}) &= -(-\mathrm{A})^\alpha \overline{f}(\overline{s}) = -\mathrm{D}^\alpha \overline{f}(\overline{s}) = -\frac{1}{\Gamma(-\alpha)} \int_0^\infty \frac{\overline{f}(\overline{s} - y) - \overline{f}(\overline{s})}{y^{\alpha+1}}\,\mathrm{d}y \\
&= -\frac{1}{\Gamma(-\alpha)} \int_0^\infty y^{-\alpha-1} (\overline{\mathcal{T}}_y - \mathbf{1}) \overline{f}(\overline{s})\,\mathrm{d}y.
\end{aligned} \tag{97}$$

PROOF. See Chapter III. □

[page 111, §1]
[111.1.1] The theorem shows that fractional derivatives of Marchaud type arise as the infinitesimal generators of coarse grained time evolutions in physics. *[111.1.2]* The order α of the derivative lies between zero and unity, and it is determined by the decay of the averaging kernel. *[111.1.3]* The order α gives a quantitative measure for the decay of the averaging kernel. *[111.1.4]* The case $\alpha \neq 1$ indicates that memory effects and history dependence may become important.

3. Applications

3.1. Fractional Invariance and Stationarity

[111.2.1] To simplify the notation $\overline{\mathrm{T}}_\alpha(\bar{t})$ will be denoted as $\mathrm{T}_\alpha(t)$ in the following. *[111.2.2]* A first application of fractional time evolutions $T_\alpha(t)$ concerns the important notion of stationarity. *[111.2.3]* This amounts to setting the left and right hand sides in eq. (2) to zero. *[111.2.4]* Surprisingly, the importance of the condition "$\mathrm{d}^\alpha f/\mathrm{d}t^\alpha$"$= 0$ for the infinitesimal generators of fractional dynamics has rarely been noticed. *[111.2.5]* Stationary states $f(s)$ may be defined more generally as states that are invariant under the time evolution after a sufficient amount of time has elapsed during which all the transients have had time to decay.

Definition 3.1 *[111.2.6] An observable or state $f(t)$ is called stationary or asymptotically invariant under the time evolution $\mathrm{T}_\alpha(t)$ if*

$$\mathrm{T}_\alpha(t)f(s) = f(s) \tag{98}$$

holds for $s/t \to \infty$. [111.2.7] It is called stationary in the strict sense, or strictly invariant under $\mathrm{T}_\alpha(t)$, if condition (98) holds for all $t \geq 0$ and $s \in \mathbb{R}$.

[111.3.1] The function $f(s) = f_0$ where f_0 is a constant is asymptotically and strictly stationary under the fractional time evolutions $\mathrm{T}_\alpha(t)$. *[111.3.2]* This follows readily by insertion into the definition, and by noting that $h_\alpha(x)$ is a probability density.

[111.4.1] In addition to the conventional constants there exists a second class of stationary states given by

$$f(s) = \begin{cases} f_0 s^{\gamma-1} & \text{for } s > 0 \\ 0 & \text{for } s \leq 0 \end{cases} \tag{99}$$

[page 112, §0] where f_0 and γ are constants. *[112.0.1]* To see this one evaluates

$$\begin{aligned}\mathrm{T}_\alpha(t)f(s) &= \int_0^\infty f(s-x)\frac{1}{t}h_\alpha\left(\frac{x}{t}\right)\,dx\\ &= f_0\int_0^s (s-x)^{\gamma-1}\frac{1}{\alpha t}H^{01}_{11}\left(\frac{x}{t}\,\middle|\,\begin{matrix}(1-1/\alpha,1/\alpha)\\(0,1)\end{matrix}\right)\mathrm{d}x\end{aligned} \tag{100}$$

where relations (170) and (172) were used to rewrite the H-function in eq. (69). *[112.0.2]* Using the integral (178), the reduction formulae (167) and (169), and property (171) one finds

$$\mathrm{T}_\alpha(t)f(s) = f_0 s^{\gamma-1}\Gamma(\gamma)H^{01}_{11}\left(\left(\frac{s}{t}\right)^\alpha\,\middle|\,\begin{matrix}(1,1)\\(1-\gamma,\alpha)\end{matrix}\right). \tag{101}$$

[112.0.3] An application of the series expansion (181) gives

$$\mathrm{T}_\alpha(t)f(s) = f_0 s^{\gamma-1}\Gamma(\gamma)\sum_{k=0}^{\infty}\frac{(-1)^k(t/s)^{k\alpha}}{k!\Gamma(\gamma-k\alpha)}. \tag{102}$$

[112.0.4] For $s/t\to\infty$ only the $k=0$ term in the series contributes and this shows that $\mathrm{T}_\alpha(t)f(s)=f(s)$ in the limit. *[112.0.5]* These considerations show that fractional time evolutions have the usual constants as strict stationary states, but admit also algebraic behaviour as a novel type of stationary states.

[112.1.1] To elucidate the significance of the new type of stationary states it is useful to consider the infinitesimal form, $\mathrm{A}_\alpha f=0$, of the stationarity condition. *[112.1.2]* The nature of the limit $s/t\to\infty$ suggests that their appearance might be related to the initial conditions. *[112.1.3]* To incorporate initial conditions into the infinitesimal generator it is necessary to consider a Riemann-Liouville representation of the fractional time derivative.

[112.2.1] The Riemann-Liouville algorithm for fractional differentiation is based on integer order derivatives of fractional integrals.

Definition 3.2 (Riemann-Liouville fractional integral) *[112.2.2] The right-sided Riemann-Liouville fractional integral of order $\alpha>0,\alpha\in\mathbb{R}$ of a locally integrable function f is defined as*

$$(\mathrm{I}^\alpha_{a+}f)(x) = \frac{1}{\Gamma(\alpha)}\int_a^x (x-y)^{\alpha-1}f(y)\,\mathrm{d}y \tag{103}$$

[page 113, §0] for $x > a$, the left-sided Riemann-Liouville fractional integral is defined as

$$(\mathrm{I}_{a-}^{\alpha}f)(x) = \frac{1}{\Gamma(\alpha)}\int_x^a (y-x)^{\alpha-1}f(y)\,\mathrm{d}y \tag{104}$$

for $x < a$.

[113.1.1] The following generalized definition, based on differentiating fractional integrals, seems to be new.

Definition 3.3 (Fractional derivatives) *[113.1.2] The (right-/left-sided) fractional derivative of order $0 < \alpha < 1$ and type $0 \le \beta \le 1$ with respect to x is defined by*

$$\mathrm{D}_{a\pm}^{\alpha,\beta}f(x) = \left(\pm\mathrm{I}_{a\pm}^{\beta(1-\alpha)}\frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{I}_{a\pm}^{(1-\beta)(1-\alpha)}f)\right)(x) \tag{105}$$

for functions for which the expression on the right hand side exists.

[113.1.3] The Riemann-Liouville fractional derivative $\mathrm{D}_{a\pm}^{\alpha} := \mathrm{D}_{a\pm}^{\alpha,0}$ corresponds to $a > -\infty$ and type $\beta = 0$. *[113.1.4]* Fractional derivatives of type $\beta = 1$ are discussed in Chapter I and were employed in [**4**]. *[113.1.5]* It seems however that fractional derivatives of general type $0 < \beta < 1$ have not been considered previously. *[113.1.6]* A relation between fractional derivatives of the same order but different types is given in Chapter IX. *[113.1.7]* For subsequent calculations it is useful to record the Laplace-Transformation

$$\mathcal{L}\left\{\mathrm{D}_{a+}^{\alpha,\beta}f(x)\right\}(u) = u^{\alpha}\mathcal{L}\left\{f(x)\right\}(u) - u^{\beta(\alpha-1)}(\mathrm{D}_{a+}^{(1-\beta)(\alpha-1),0}f)(0+) \tag{106}$$

where the inital value $(\mathrm{D}_{a+}^{(1-\beta)(\alpha-1),0}f)(0+)$ is the Riemann-Liouville derivative for $t \to 0+$. *[113.1.8]* Note that fractional derivatives of type 1 involve nonfractional initial values.

[113.2.1] It is now possible to discuss the infinitesimal form of fractional stationarity where the generator A_{α} for initial conditions of type $0 \le \beta \le 1$ is represented by $\mathrm{D}_{0+}^{\alpha,\beta}$. *[113.2.2]* The fractional differential equation

$$\mathrm{D}_{0+}^{\alpha,\beta}f(t) = 0 \tag{107}$$

for f with initial condition

$$\mathrm{I}_{0+}^{(1-\beta)(1-\alpha)}f(0+) = f_0 \tag{108}$$

[page 114, §0] defines fractional stationarity of order α and type β. *[114.0.1]* Of course, for $\alpha = 1$ this definition reduces to the conventional definition of stationarity. *[114.0.2]* Equation (107) is solved by

$$f(t) = \frac{f_0\, t^{(1-\beta)(\alpha-1)}}{\Gamma((1-\beta)(\alpha-1)+1)}. \tag{109}$$

[114.0.3] This may be seen by inserting $f(t)$ into the definition

$$\mathrm{D}_{0+}^{\alpha,\beta} f(x) = \left(\mathrm{I}_{0+}^{\beta(1-\alpha)} \frac{\mathrm{d}}{\mathrm{d}x} (\mathrm{I}_{0+}^{(1-\beta)(1-\alpha)} f)\right)(x) \tag{110}$$

and using the basic fractional integral

$$\mathrm{I}_{a+}^{\alpha}(x-a)^{\beta} = \frac{\Gamma(\beta+1)}{\Gamma(\alpha+\beta+1)}(x-a)^{\alpha+\beta} \tag{111}$$

(derived in eq. (1.30) in Chapter I). *[114.0.4]* Note that the fractional integral

$$\mathrm{I}_{0+}^{(1-\beta)(1-\alpha)} f(t) = f_0 \tag{112}$$

remains conserved and constant for all t while the function itself varies. *[114.0.5]* In particular $\lim_{t\to 0} f(t) = \infty$ and $\lim_{t\to\infty} f(t) = 0$. *[114.0.6]* For $\beta = 1$ and for $\alpha = 1$ one recovers $f(t) = f_0$ as usual.

[114.1.1] The new types of stationary states for which a fractional integral rather than the function itself is constant were first discussed in [**6, 9**]. *[114.1.2]* It seems to me that the lack of knowledge about fractional stationarity is partially responsible for the difficulty of deciding which type of fractional derivative should be used when generalizing traditional equations of motion.

[114.2.1] Another simple instance of a fractional differential equation is the equation

$$\mathrm{D}_{0+}^{\alpha,\beta} f(t) = C \tag{113}$$

with $C \in \mathbb{R}$ a constant, and with initial condition

$$\mathrm{I}_{0+}^{(1-\beta)(1-\alpha)} f(0+) = f_0 \tag{114}$$

as before. *[114.2.2]* Laplace transformation using eq. (106) gives

$$f(u) = \frac{C}{u^{\alpha+1}} + \frac{f_0}{u^{\alpha+\beta(1-\alpha)}} \tag{115}$$

and thence

$$f(t) = \frac{C\, t^{\alpha}}{\Gamma(\alpha+1)} + \frac{f_0\, t^{(1-\beta)(\alpha-1)}}{\Gamma((1-\beta)(1-\alpha)+1)}. \tag{116}$$

[page 115, §0] *[115.0.1]* For $\beta = 1$ this reduces to

$$f(t) = \frac{C\ t^{\alpha}}{\Gamma(\alpha+1)} + f_0. \tag{117}$$

3.2. Generalized Fractional Relaxation

[115.1.1] Consider the fractional Cauchy problem

$$\mathrm{D}^{\alpha,\beta}_{0+} f(t) = -C\ f(t) \tag{118}$$

for f with initial condition

$$\mathrm{I}^{(1-\beta)(1-\alpha)}_{0+} f(0+) = f_0 \tag{119}$$

where C is a ("fractional relaxation") constant. *[115.1.2]* Laplace Transformation gives

$$f(u) = \frac{u^{\beta(\alpha-1)}\ f_0}{C + u^{\alpha}}. \tag{120}$$

[115.1.3] To invert the Laplace transform rewrite this equation as

$$f(u) = \frac{u^{\alpha-\gamma}}{C+u^{\alpha}} = u^{-\gamma}\frac{1}{Cu^{-\alpha}+1} = \sum_{k=0}^{\infty}(-C)^k u^{-\alpha k-\gamma} \tag{121}$$

with

$$\gamma = \alpha + \beta(1-\alpha). \tag{122}$$

[115.1.4] Inverting the series term by term using $\mathcal{L}\left\{x^{\alpha-1}/\Gamma(\alpha)\right\} = u^{-\alpha}$ yields the result

$$f(t) = t^{\gamma-1}\sum_{k=0}^{\infty}\frac{(-Ct^{\alpha})^k}{\Gamma(\alpha k+\gamma)}\ . \tag{123}$$

[115.1.5] The solution may be written as

$$f(t) = f_0\ t^{(1-\beta)(\alpha-1)} E_{\alpha,\alpha+\beta(1-\alpha)}(-Ct^{\alpha}) \tag{124}$$

using the generalized Mittag-Leffler function defined by

$$E_{a,b}(x) = \sum_{k=0}^{\infty}\frac{x^k}{\Gamma(ak+b)} \tag{125}$$

[page 116, §0] for all $a > 0, b \in \mathbb{C}$. *[116.0.1]* This function is an entire function of order $1/a$ [**38**]. *[116.0.2]* Moreover it is completely monotone if and only if $0 < a \leq 1$ and $b \geq a$ [**39**].

[116.1.1] For $C = 0$ the result reduces to eq. (109) because $E_{a,b}(0) = 1/\Gamma(b)$. *[116.1.2]* Of special interest is again the case $\beta = 1$. *[116.1.3]* It has the well known solution

$$f(t) = f_0 \; E_\alpha(-Ct^\alpha) \tag{126}$$

where $E_\alpha(x) = E_{\alpha,1}(x)$ denotes the ordinary Mittag-Leffler function.

3.3. Generalized Fractional Diffusion

[116.2.1] Consider the fractional partial differential equation for $f : \mathbb{R}^d \times \mathbb{R}_+ \to \mathbb{R}$

$$\mathrm{D}^{\alpha,\beta}_{0+} f(\boldsymbol{r}, t) = C \; \Delta f(\boldsymbol{r}, t) \tag{127}$$

with Laplacian Δ and fractional "diffusion" constant C. *[116.2.2]* The function $f(\boldsymbol{r}, t)$ is assumed to obey the initial condition

$$\mathrm{I}^{(1-\beta)(1-\alpha)}_{0+} f(\boldsymbol{r}, 0+) = f_{0\boldsymbol{r}} = f_0 \delta(\boldsymbol{r}) \tag{128}$$

where $\delta(\boldsymbol{r})$ is the Dirac measure at the origin. *[116.2.3]* Fourier Transformation, defined as

$$\mathcal{F}\left\{f(\boldsymbol{r})\right\}(\boldsymbol{q}) = \int_{\mathbb{R}^d} e^{i\boldsymbol{q}\cdot\boldsymbol{r}} f(\boldsymbol{r}) \mathrm{d}\boldsymbol{r}, \tag{129}$$

and Laplace transformation of eq. (127) now yields

$$f(\boldsymbol{q}, u) = \frac{u^{\beta(\alpha-1)} \; f_0}{C\boldsymbol{q}^2 + u^\alpha}. \tag{130}$$

[116.2.4] Using the result (124) for the inverse Laplace transform of (120) gives

$$f(\boldsymbol{q}, t) = f_0 \; t^{(1-\beta)(\alpha-1)} E_{\alpha,\alpha+\beta(1-\alpha)}(-C\boldsymbol{q}^2 t^\alpha). \tag{131}$$

[116.2.5] Setting $\boldsymbol{q} = 0$ shows that the solution of (127) cannot be a probability density except for $\beta = 1$. *[116.2.6]* For $\beta \neq 1$ the spatial integral is time dependent, and f would need to be divided by $t^{(1-\beta)(\alpha-1)}$ to admit a probabilistic interpretation.

[116.3.1] To invert eq. (130) completely it seems advantageous to first invert the Fourier transform and then the Laplace transform. *[116.3.2]* The Fourier transform may be inverted by noting the formula [**40**]

$$(2\pi)^{-d/2} \int e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \left(\frac{|\boldsymbol{r}|}{m}\right)^{1-(d/2)} K_{(d-2)/2}\left(m|\boldsymbol{r}|\right) \; \mathrm{d}\boldsymbol{r} = \frac{1}{\boldsymbol{q}^2 + m^2} \tag{132}$$

[page 117, §0] which leads to

$$f(\boldsymbol{r},u) = f_0(2\pi C)^{-d/2}\left(\frac{r}{\sqrt{C}}\right)^{1-(d/2)} u^{\beta(\alpha-1)+\alpha(d-2)/4}K_{(d-2)/2}\left(\frac{ru^{\alpha/2}}{\sqrt{C}}\right) \quad (133)$$

with $r = |\boldsymbol{r}|$. *[117.0.1]* To invert the Laplace transform one uses again the relation (78) with the Mellin transform defined in eq. (79). *[117.0.2]* Setting $A = r/\sqrt{C}$, $\lambda = \alpha/2$, $\nu = (d-2)/2$ and $\mu = \beta(\alpha-1) + \alpha(d-2)/4$ and using the general relation

$$\mathcal{M}\left\{x^q g(bx^p)\right\}(s) = \frac{1}{p}b^{-(s+q)/p}g\left(\frac{s+q}{p}\right) \qquad (b,p>0) \quad (134)$$

leads to

$$\mathcal{M}\left\{f(r,u)\right\}(s) = \frac{f_0}{\lambda}(2\pi C)^{-d/2}A^{1-(d/2)}A^{-(s+\mu)/\lambda}\mathcal{M}\left\{K_\nu(u)\right\}((s+\mu)/\lambda)\,. \quad (135)$$

[117.0.3] The Mellin transform of the Bessel function reads [**32**]

$$\mathcal{M}\left\{K_\nu(x)\right\}(s) = 2^{s-2}\Gamma\left(\frac{s+\nu}{2}\right)\Gamma\left(\frac{s-\nu}{2}\right). \quad (136)$$

[117.0.4] Inserting this, using eq.(78), and restoring the original variables then yields

$$\mathcal{M}\left\{f(r,t)\right\}(s) = \frac{f_0}{\alpha(r^2\pi)^{d/2}}\left(\frac{r}{2\sqrt{C}}\right)^{2(1-\beta)(1-(1/\alpha))}\left(\frac{r}{2\sqrt{C}}\right)^{2s/\alpha} \frac{\Gamma\left(\frac{d}{2}+(\beta-1)(1-\frac{1}{\alpha})-\frac{s}{\alpha}\right)\Gamma\left(1+(\beta-1)(1-\frac{1}{\alpha})-\frac{s}{\alpha}\right)}{\Gamma(1-s)} \quad (137)$$

for the Mellin transform of f. *[117.0.5]* Comparing this with the Mellin transform of the H-function in eq. (175) allows to identify the H-function parameters as $m = 0$,$n = 2$,$p = 2$,$q = 1$, $A_1 = A_2 = 1/\alpha$, $a_1 = 1-(d/2)-(\beta-1)(1-(1/\alpha))$, $a_2 = (1-\beta)(1-(1/\alpha))$, $b_1 = 0$ and $B_1 = 1$ if $(\alpha d/2) + (\beta-1)(\alpha-1) > 0$. *[117.0.6]* Then the result becomes

$$f(r,t) = \frac{f_0}{\alpha(r^2\pi)^{d/2}}\left(\frac{r}{2\sqrt{C}}\right)^{2(1-\beta)(1-(1/\alpha))} H^{02}_{21}\left(\left(\frac{2\sqrt{C}}{r}\right)^{2/\alpha}t\,\middle|\,\begin{matrix}(1-\frac{d}{2}+(1-\beta)(1-\frac{1}{\alpha}),\frac{1}{\alpha}),((1-\beta)(1-\frac{1}{\alpha}),\frac{1}{\alpha})\\(0,1)\end{matrix}\right). \quad (138)$$

[page 118, §0] *[118.0.1]* This may be simplified using eqs.(170), (171) and (172) to become finally

$$f(r,t) = \frac{f_0\ t^{(1-\beta)(\alpha-1)}}{(r^2\pi)^{d/2}} H^{20}_{12}\left(\frac{r^2}{4Ct^\alpha}\ \middle|\ \begin{matrix}(1+(1-\beta)(\alpha-1),\alpha)\\ (d/2,1),(1,1)\end{matrix}\right). \tag{139}$$

[118.0.2] The result reduces to the known result [**15, 8**] for $\beta = 1$. *[118.0.3]* In that case $f(\boldsymbol{r},t)$ is also a probability density. *[118.0.4]* For $\beta \neq 1$ the function $f(\boldsymbol{r},t)$ does not have a probabilistic interpretation because its normalization decays as $t^{(1-\beta)(\alpha-1)}$.

3.4. Relation with Continuous Time Random Walk

[118.1.1] The fractional diffusion eq. (127) of type $\beta = 1$ has a probabilistic interpretation as noted after eq. (131). *[118.1.2]* $f(\boldsymbol{r},t)$ may be viewed as the probability density for a random walker or diffusing object to be at position $\boldsymbol{r}$ at time t under the condition that it started from the origin $\boldsymbol{r} = \boldsymbol{0}$ at time $t = 0$. This probabilistic interpretation is very helpful for understanding the meaning of the fractional time derivative appearing in eq. (127). *[118.1.3]* Rewriting equation (127) in integral form it becomes

$$f(\boldsymbol{r},t) = \delta_{\boldsymbol{r}\boldsymbol{0}} + \frac{C}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\Delta f(\boldsymbol{r},t)\ \mathrm{d}s \tag{140}$$

where the initial condition has been incorporated. *[118.1.4]* This integral equation is very reminiscent of the integral equation for continuous time random walks [**41, 42**].

[118.2.1] In a continuous time random walk one imagines a random walker that starts at $\boldsymbol{r} = \boldsymbol{0}$ at time $t = 0$ and proceeds by successive random jumps [**43, 44, 45, 46, 47, 48**]. *[118.2.2]* The probability density for a time interval of length t between two consecutive jumps is denoted $\psi(t)$ and the probability density of a displacement by a vector $\boldsymbol{r}$ in a single jump is denoted $p(\boldsymbol{r})$. *[118.2.3]* Then the integral equation of continuous time random walk theory reads

$$f(\boldsymbol{r},t) = \delta_{\boldsymbol{r}\boldsymbol{0}}\Phi(t) + \int_0^t \psi(t-s)\int_{\mathbb{R}^d} p(\boldsymbol{r}-\boldsymbol{r}')f(\boldsymbol{r},t)\ \mathrm{d}\boldsymbol{r}'\ \mathrm{d}s \tag{141}$$

where $\Phi(t)$ is the probability that the walker survives at the origin for a time of length t. *[118.2.4]* Here the walker is assumed to be prepared in its initial position from which it develops according to $\psi(t)$. *[118.2.5]* In general the first step needs special consideration [**49, 49, 45**]. *[118.2.6]* The survival probablity $\Phi(t)$ is related to the waiting *[page 119, §0]* time density through

$$\Phi(t) = 1 - \int_0^t \psi(t')\, \mathrm{d}t'. \tag{142}$$

[119.1.1] The formal similarity between eqs. (141) and (140) suggests that there exists a relation between them. *[119.1.2]* To establish the relation note that eq. (130) for $\beta = 1$ gives the solution of eq. (127) in Fourier-Laplace space as

$$f(\boldsymbol{q}, u) = \frac{u^{\alpha-1}}{C\boldsymbol{q}^2 + u^\alpha}. \tag{143}$$

[119.1.3] The Fourier-Laplace solution of eq.(141) is [**44, 50, 51, 46**]

$$f(\boldsymbol{q}, u) = \frac{1}{u}\frac{1 - \psi(u)}{1 - \psi(u)p(\boldsymbol{q})}. \tag{144}$$

[119.1.4] Equating these two equations yields

$$\frac{1 - p(\boldsymbol{q})}{C\boldsymbol{q}^2} = \frac{1 - \psi(u)}{u^\alpha \psi(u)}. \tag{145}$$

[119.1.5] Because the left hand side does not depend on u and the right hand side is independent of $\boldsymbol{q}$ they must both equal a common constant τ_0^α. *[119.1.6]* It follows that

$$p(\boldsymbol{q}) = 1 - C\tau_0^\alpha \boldsymbol{q}^2 \tag{146}$$

identifying the constant $C\tau_0^\alpha$ as the mean square displacement of a single jump. *[119.1.7]* For the waiting time density one finds

$$\psi(u) = \frac{1}{1 + \tau_0^\alpha u^\alpha}, \tag{147}$$

which may be inverted in the same way as eq. (120) to give

$$\psi(t; \alpha, \tau_0) = \frac{1}{\tau_0}\left(\frac{t}{\tau_0}\right)^{\alpha-1} E_{\alpha,\alpha}\left(-\frac{t^\alpha}{\tau_0^\alpha}\right) \tag{148}$$

where $E_{a,b}(x)$ is again the Mittag-Leffler function defined in eq. (40).

[119.2.1] For $\alpha = 1$ the waiting time density becomes exponential

$$\psi(t; 1, \tau_0) = \frac{1}{\tau_0} e^{-t/\tau_0}. \tag{149}$$

[page 120, §0] *[120.0.1]* For $0 < \alpha < 1$ characteristic differences arise from the asymptotic behaviour for $t \to 0$ and $t \to \infty$. *[120.0.2]* The asymptotic behaviour of $\psi(t)$ for $t \to 0$ is obtained by noting that $E_{\alpha,\alpha}(0) = 1$, and hence

$$\psi(t) \sim t^{\alpha-1} \tag{150}$$

for $t \to 0$. *[120.0.3]* For $\alpha < 1$ the waiting time density is singular at the origin implying a statistical abundance of short intervals between jumps compared to the exponential case $\alpha = 1$. *[120.0.4]* For large $t \to \infty$ recall the asymptotic series expansion [**52**]

$$E_{a,b}(z) = -\sum_{n=1}^{N} \frac{z^{-n}}{\Gamma(b - an)} + O(|z|^N) \tag{151}$$

valid for $|\arg(-z)| < (1 - (a/2))\pi$ and $z \to \infty$. *[120.0.5]* It follows that $E_{a,a}(-x) \sim x^{-2}$ for $x \to \infty$ and hence

$$\psi(t) \sim t^{-1-\alpha} \tag{152}$$

for $t \to \infty$. *[120.0.6]* This shows that fractional diffusion is equivalent to a continuous time random walk whose waiting time density is a generalized Mittag-Leffler function. *[120.0.7]* The waiting time density has a long time tail of the form usually assumed in the general theory [**53, 49, 54, 46**] and exhibits a power law divergence at the origin. *[120.0.8]* The exponent of both power laws is given by the order of the fractional derivative.

4. H-Functions

4.1. Definition

[120.1.1] The H-function of order $(m, n, p, q) \in \mathbb{N}^4$ and with parameters $A_i \in \mathbb{R}_+ (i = 1, \dots, p)$, $B_i \in \mathbb{R}_+ (i = 1, \dots, q)$, $a_i \in \mathbb{C} (i = 1, \dots, p)$, and $b_i \in \mathbb{C} (i = 1, \dots, q)$ is defined for $z \in \mathbb{C}, z \neq 0$ by the contour integral [**55, 56, 57, 58, 59**]

$$H^{m,n}_{p,q}\left(z \,\middle|\, \begin{matrix} (a_1, A_1), \dots, (a_p, A_p) \\ (b_1, B_1), \dots, (b_q, B_q) \end{matrix} \right) = \frac{1}{2\pi i} \int_{\mathcal{L}} \eta(s) z^{-s} \, \mathrm{d}s \tag{153}$$

where the integrand is

$$\eta(s) = \frac{\prod_{i=1}^{m} \Gamma(b_i + B_i s) \prod_{i=1}^{n} \Gamma(1 - a_i - A_i s)}{\prod_{i=n+1}^{p} \Gamma(a_i + A_i s) \prod_{i=m+1}^{q} \Gamma(1 - b_i - B_i s)}. \tag{154}$$

[page 121, §0] *[121.0.1]* In (153) $z^{-s} = \exp\{-s\log|z| - i\arg z\}$ and $\arg z$ is not necessarily the principal value. *[121.0.2]* The integers m, n, p, q must satisfy

$$0 \le m \le q, \qquad 0 \le n \le p, \tag{155}$$

and empty products are interpreted as being unity. *[121.0.3]* The parameters are restricted by the condition

$$\mathbb{P}_a \cap \mathbb{P}_b = \emptyset \tag{156}$$

where

$$\begin{aligned}\mathbb{P}_a &= \{\text{poles of } \Gamma(1 - a_i - A_i s)\} = \left\{\frac{1 - a_i + k}{A_i} \in \mathbb{C} : i = 1, \ldots, n; k \in \mathbb{N}_0\right\} \\ \mathbb{P}_b &= \{\text{poles of } \Gamma(b_i + B_i s)\} = \left\{\frac{-b_i - k}{B_i} \in \mathbb{C} : i = 1, \ldots, m; k \in \mathbb{N}_0\right\}\end{aligned} \tag{157}$$

are the poles of the numerator in (154). *[121.0.4]* The integral converges if one of the following conditions holds [**59**]

$$\mathcal{L} = \mathcal{L}(c - i\infty, c + i\infty; \mathbb{P}_a, \mathbb{P}_b); \quad |\arg z| < C\pi/2; \quad C > 0 \tag{158a}$$

$$\mathcal{L} = \mathcal{L}(c - i\infty, c + i\infty; \mathbb{P}_a, \mathbb{P}_b); \quad |\arg z| = C\pi/2; \quad C \ge 0; \quad cD < -\operatorname{Re} F \tag{158b}$$

$$\mathcal{L} = \mathcal{L}(-\infty + i\gamma_1, -\infty + i\gamma_2; \mathbb{P}_a, \mathbb{P}_b); \quad D > 0; \quad 0 < |z| < \infty \tag{159a}$$

$$\mathcal{L} = \mathcal{L}(-\infty + i\gamma_1, -\infty + i\gamma_2; \mathbb{P}_a, \mathbb{P}_b); \quad D = 0; \quad 0 < |z| < E^{-1} \tag{159b}$$

$$\mathcal{L} = \mathcal{L}(-\infty + i\gamma_1, -\infty + i\gamma_2; \mathbb{P}_a, \mathbb{P}_b); \quad D = 0; \quad |z| = E^{-1}; C \ge 0; \operatorname{Re} F < 0 \tag{159c}$$

$$\mathcal{L} = \mathcal{L}(\infty + i\gamma_1, \infty + i\gamma_2; \mathbb{P}_a, \mathbb{P}_b); \quad D < 0; \quad 0 < |z| < \infty \tag{160a}$$

$$\mathcal{L} = \mathcal{L}(\infty + i\gamma_1, \infty + i\gamma_2; \mathbb{P}_a, \mathbb{P}_b); \quad D = 0; \quad |z| > E^{-1} \tag{160b}$$

$$\mathcal{L} = \mathcal{L}(\infty + i\gamma_1, \infty + i\gamma_2; \mathbb{P}_a, \mathbb{P}_b); \quad D = 0; \quad |z| = E^{-1}; C \ge 0; \operatorname{Re} F < 0 \tag{160c}$$

where $\gamma_1 < \gamma_2$. *[121.0.5]* Here $\mathcal{L}(z_1, z_2; \mathbb{G}_1, \mathbb{G}_2)$ denotes a contour in the complex plane starting at z_1 and ending at z_2 and separating the points in $\mathbb{G}_1$ from those *[page 122, §0]* in $\mathbb{G}_2$, and the notation

$$C = \sum_{i=1}^{n} A_i - \sum_{i=n+1}^{p} A_i + \sum_{i=1}^{m} B_i - \sum_{i=m+1}^{q} B_i \tag{161}$$

$$D = \sum_{i=1}^{q} B_i - \sum_{i=1}^{p} A_i \tag{162}$$

$$E = \prod_{i=1}^{p} A_i^{A_i} \prod_{i=1}^{q} B_i^{-B_i} \tag{163}$$

$$F = \sum_{i=1}^{q} b_i - \sum_{i=1}^{p} a_j + (p-q)/2 + 1 \tag{164}$$

was employed. *[122.0.1]* The H-functions are analytic for $z \neq 0$ and multivalued (single valued on the Riemann surface of $\log z$).

4.2. Basic Properties

[122.1.1] From the definition of the H-functions follow some basic properties. *[122.1.2]* Let $S_n (n \geq 1)$ denote the symmetric group of n elements, and let π_n denote a permutation in S_n. *[122.1.3]* Then the product structure of (154) implies that for all $\pi_n \in S_n, \pi_m \in S_m, \pi_{p-n} \in S_{p-n}$ and $\pi_{q-m} \in S_{q-m}$

$$H^{m,n}_{p,q}\left(z \,\middle|\, \begin{matrix} (a_1, A_1), \dots, (a_p, A_p) \\ (b_1, B_1), \dots, (b_q, B_q) \end{matrix}\right) = H^{m,n}_{p,q}\left(z \,\middle|\, \begin{matrix} P_n, P_{p-n} \\ P_m, P_{q-m} \end{matrix}\right) \tag{165}$$

where the parameter permutations

$$\begin{aligned} P_n &= (a_{\pi_n(1)}, A_{\pi_n(1)}), \dots, (a_{\pi_n(n)}, A_{\pi_n(n)}) \\ P_{p-n} &= (a_{\pi_{p-n}(n+1)}, A_{\pi_{p-n}(n+1)}), \dots, (a_{\pi_{p-n}(p)}, A_{\pi_{p-n}(p)}) \\ P_m &= (b_{\pi_m(1)}, B_{\pi_m(1)}), \dots, (b_{\pi_m(m)}, B_{\pi_m(m)}) \\ P_{q-m} &= (b_{\pi_{q-m}(m+1)}, B_{\pi_{q-m}(m+1)}), \dots, (b_{\pi_{q-m}(q)}, B_{\pi_{q-m}(q)}) \end{aligned} \tag{166}$$

have to be inserted on the right hand side. *[122.1.4]* If any of $n, m, p-n$ or $q-m$ vanishes the corresponding permutation is absent.

[122.2.1] The order reduction formula

$$\begin{aligned} &H^{m,n}_{p,q}\left(z \,\middle|\, \begin{matrix} (a_1, A_1), (a_2, A_2) \dots, (a_p, A_p) \\ (b_1, B_1), (b_2, B_2) \dots, (b_{q-1}, B_{q-1})(a_1, A_1) \end{matrix}\right) \\ &= H^{m,n-1}_{p-1,q-1}\left(z \,\middle|\, \begin{matrix} (a_2, A_2), \dots, (a_p, A_p) \\ (b_1, B_1), \dots, (b_{q-1}, B_{q-1}) \end{matrix}\right) \end{aligned} \tag{167}$$

[page 123, §0] holds for $n \geq 1$ and $q > m$, and similarly

$$H^{m,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a_1,A_1),(a_2,A_2)\dots,(a_{p-1},A_{p-1})(b_1,B_1)\\(b_1,B_1),(b_2,B_2)\dots,(b_q,B_q)\end{array}\right)$$
$$= H^{m-1,n}_{p-1,q-1}\left(z \,\middle|\, \begin{array}{l}(a_1,A_1),\dots,(a_{p-1},A_{p-1})\\(b_2,B_2),\dots,(b_q,B_q)\end{array}\right) \tag{168}$$

for $m \geq 1$ and $p > n$. *[123.0.1]* The formula

$$H^{m,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a,0),(a_2,A_2)\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right)$$
$$= \Gamma(1-a)H^{m,n-1}_{p-1,q}\left(z \,\middle|\, \begin{array}{l}(a_2,A_2),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right) \tag{169}$$

holds for $n \geq 1$ and $\mathrm{Re}\,(1-a) > 0$. *[123.0.2]* Analogous formulae are readily found if a parameter pair $(a,0)$ or $(b,0)$ appears in one of the other groups.

[123.1.1] A change of variables in (153) shows

$$H^{m,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a_1,A_1),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right)$$
$$= H^{n,m}_{q,p}\left(\frac{1}{z} \,\middle|\, \begin{array}{l}(1-b_1,B_1),\dots,(1-b_q,B_q)\\(1-a_1,A_1),\dots,(1-a_p,A_p)\end{array}\right) \tag{170}$$

which allows to transform an H-function with $D > 0$ and $\arg z$ to one with $D < 0$ and $\arg(1/z)$. *[123.1.2]* For $\gamma > 0$

$$\frac{1}{\gamma}H^{m,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a_1,A_1),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right)$$
$$= H^{m,n}_{p,q}\left(z^\gamma \,\middle|\, \begin{array}{l}(a_1,\gamma A_1),\dots,(a_p,\gamma A_p)\\(b_1,\gamma B_1),\dots,(b_q,\gamma B_q)\end{array}\right) \tag{171}$$

while for $\gamma \in \mathbb{R}$

$$z^\gamma H^{m,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a_1,A_1),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right)$$
$$= H^{m,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a_1+\gamma A_1,A_1),\dots,(a_p+\gamma A_p,A_p)\\(b_1+\gamma B_1,B_1),\dots,(b_q+\gamma B_q,B_q)\end{array}\right) \tag{172}$$

[page 124, §0] holds.

[124.1.1] For $m = 0$ with conditions (159) the integrand is analytic and thus

$$H^{0,n}_{p,q}\left(z \,\middle|\, \begin{array}{l}(a_1,A_1),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right) = 0. \tag{173}$$

4.3. Integral Transformations

[124.2.1] The definition of an H-function in eq. (153) becomes an inverse Mellin transform if $\mathcal{L}$ is chosen parallel to the imaginary axis inside the strip

$$\max_{1\le i\le m} \operatorname{Re}\frac{-b_i}{B_i} < s < \min_{1\le i\le m} \operatorname{Re}\frac{1-a_i}{A_i} \tag{174}$$

by the Mellin inversion theorem [**60**]. *[124.2.2]* Therefore

$$\mathcal{M}\left\{H^{m,n}_{p,q}(z)\right\}(s) = \eta(s) = \frac{\prod_{i=1}^{m}\Gamma(b_i+B_is)\prod_{i=1}^{n}\Gamma(1-a_i-A_is)}{\prod_{i=n+1}^{p}\Gamma(a_i+A_is)\prod_{i=m+1}^{q}\Gamma(1-b_i-B_is)} \tag{175}$$

whenever the inequality

$$\max_{1\le i\le m} \operatorname{Re}\frac{-b_i}{B_i} < \min_{1\le i\le m} \operatorname{Re}\frac{1-a_i}{A_i} \tag{176}$$

is fulfilled.

[124.3.1] The Laplace transform of an H-function is obtained from eq. (175) by using eq. (78). *[124.3.2]* One finds

$$\begin{aligned}\mathcal{L}\left\{H^{m,n}_{p,q}(z)\right\}(u) &= \int_0^\infty e^{-ux} H^{m,n}_{p,q}\left(x \,\middle|\, \begin{array}{l}(a_1,A_1),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right)\mathrm{d}x\\ &= H^{n+1,m}_{q,p+1}\left(u \,\middle|\, \begin{array}{l}(1-b_1-B_1,B_1),\dots,(1-b_q-B_q,B_q)\\(0,1)(1-a_1-A_1,A_1),\dots,(1-a_p-A_p,A_p)\end{array}\right)\\ &= \frac{1}{u}H^{m,n+1}_{p+1,q}\left(\frac{1}{u} \,\middle|\, \begin{array}{l}(0,1)(a_1,A_1),\dots,(a_p,A_p)\\(b_1,B_1),\dots,(b_q,B_q)\end{array}\right)\end{aligned} \tag{177}$$

valid for $\operatorname{Re} s > 0$, $C > 0$, $|\arg z| < \frac{1}{2}C\pi$ and $\min_{1\le j\le m}\operatorname{Re}(b_j/B_j) > -1$.

[page 125, §1]

[125.1.1] The definite integral found in [**59**, 2.25.2.2.]

$$\int_0^y x^{\beta-1}(y-x)^{\gamma-1} H_{p,q}^{m,n}\left(Cx^\delta(y-x)^\eta \,\middle|\, \begin{matrix}(a_1,A_1),\ldots,(a_p,A_p)\\(b_1,B_1),\ldots,(b_q,B_q)\end{matrix}\right)\mathrm{d}x$$
$$= y^{\beta+\gamma-1} H_{p+2,q+1}^{m,n+2}\left(Cy^{\delta+\eta} \,\middle|\, \begin{matrix}(1-\beta,\delta),(1-\gamma,\eta),(a_1,A_1),\ldots,(a_p,A_p)\\(b_1,B_1),\ldots,(b_q,B_q),(1-\beta-\gamma,\delta+\eta)\end{matrix}\right) \quad (178)$$

contains as a special case the fractional Riemann-Liouville integral [**58**, (2.7.13)]

$$\mathrm{I}_{0+}^\alpha H_{p,q}^{m,n}(y) = \frac{1}{\Gamma(\alpha)}\int_0^y (y-x)^{\alpha-1} H_{p,q}^{m,n}\left(x \,\middle|\, \begin{matrix}(a_1,A_1),\ldots,(a_p,A_p)\\(b_1,B_1),\ldots,(b_q,B_q)\end{matrix}\right)\mathrm{d}x$$
$$= y^\alpha H_{p+1,q+1}^{m,n+1}\left(y \,\middle|\, \begin{matrix}(0,1),(a_1,A_1),\ldots,(a_p,A_p)\\(b_1,B_1),\ldots,(b_q,B_q),(-\alpha,1)\end{matrix}\right) \quad (179)$$

valid if $\min_{1\le j\le m} \mathrm{Re}\,(b_j/B_j) > 0$. *[125.1.2]* The fractional Riemann-Liouville derivative is obtained from this formula by analytic continuation to $\alpha < 0$.

4.4. Series Expansions

[125.2.1] The H-functions may be represented as the series [**56, 57, 58, 59**]

$$H_{p,q}^{m,n}\left(z \,\middle|\, \begin{matrix}(a_1,A_1),\ldots,(a_p,A_p)\\(b_1,B_1),\ldots,(b_q,B_q)\end{matrix}\right) = \sum_{i=1}^m\sum_{k=0}^\infty c_{ik}\frac{(-1)^k}{k!B_i} z^{(b_i+k)/B_i} \quad (180a)$$

where

$$c_{ik} = \frac{\displaystyle\prod_{\substack{j=1\\j\ne i}}^m \Gamma(b_j-(b_i+k)B_j/B_i)\prod_{j=1}^n \Gamma(1-a_j+(b_i+k)A_j/B_i)}{\displaystyle\prod_{j=m+1}^q \Gamma(1-b_j+(b_i+k)B_j/B_i)\prod_{j=n+1}^p \Gamma(a_j-(b_i+k)A_j/B_i)} \quad (180b)$$

whenever $D \ge 0$, $\mathcal{L}$ is as in (158) or (159) and the poles in $\mathbb{P}_b$ are simple. *[125.2.2]* Similarly

$$H_{p,q}^{m,n}\left(z \,\middle|\, \begin{matrix}(a_1,A_1),\ldots,(a_p,A_p)\\(b_1,B_1),\ldots,(b_q,B_q)\end{matrix}\right) = \sum_{i=1}^n\sum_{k=0}^\infty c_{ik}\frac{(-1)^k}{k!A_i} z^{-(1-a_i+k)/A_i} \quad (181a)$$

[page 126, §0] where

$$c_{ik} = \frac{\prod\limits_{\substack{j=1\\ j\neq i}}^{n} \Gamma(1-a_j-(1-a_i+k)A_j/A_i) \prod\limits_{j=1}^{m} \Gamma(b_j+(1-a_i+k)B_j/A_i)}{\prod\limits_{j=n+1}^{p} \Gamma(a_j+(1-a_i+k)A_j/A_i) \prod\limits_{j=m+1}^{q} \Gamma(1-b_j-(1-a_i+k)B_j/A_i)} \tag{181b}$$

whenever $D \leq 0$, $\mathcal{L}$ is as in (158) or (160) and the poles in $\mathbb{P}_a$ are simple.

5. Appendix: Proof of Proposition 2.2

[126.1.1] The proof given below follows Ref. [**30**]. *[126.1.2]* Suppose $\lim_{n\to\infty} \mu_n(s) = \mu(s)$ and $\lim_{n\to\infty} \mu_n(a_n s + b_n) = \nu(s)$ with $\mu(s)$ and $\nu(s)$ both nondegenerate. *[126.1.3]* Then it must be shown that there exist $a > 0$ and b such that

$$\mu(s) = \nu(as + b). \tag{182}$$

[126.1.4] Pick a sequence of integers $n_1 < n_2 < \ldots < n_k < \ldots$ such that $\lim_{k\to\infty} a_{n_k} = a$ and $\lim_{k\to\infty} b_{n_k} = b$ exist with $0 \leq a \leq \infty$ and $-\infty \leq b \leq \infty$. *[126.1.5]* Consider this sequence of indices from now on as fixed. *[126.1.6]* Then, to simplify the notation, suppose without loss of generality that $\lim_{k\to\infty} a_k = a$ and $\lim_{k\to\infty} b_k = b$.

[126.2.1] First it will be shown that $0 < a < \infty$. *[126.2.2]* Suppose $a = \infty$. *[126.2.3]* Let

$$u = \sup\{x : \limsup_{n\to\infty}(a_n x + b_n) < \infty\}. \tag{183}$$

[126.2.4] Then for $v < x < u$

$$\limsup_{n\to\infty}(a_n v + b_n) \leq \limsup_{n\to\infty}(v - x)a_n + \limsup_{n\to\infty}(a_n x + b_n), \tag{184}$$

and hence for every $v < u$ it follows that $\nu(v) = 0$ because $(a_n v + b_n) \to -\infty$ with $n \to \infty$. *[126.2.5]* For $v > u$, on the other hand, $\limsup(a_n v + b_n) = \infty$ and hence $\nu(v) = 1$ for $v > u$. *[126.2.6]* Thus the assumption $a = \infty$ contradicts to $\nu(s)$ being nondegenerate.

[126.3.1] It follows that also b must be finite. *[126.3.2]* In fact if $\lim_{n\to\infty}(a_n x + b_n) = \infty$ then $\nu(x) = 1$ while for $\lim_{n\to\infty}(a_n x + b_n) = -\infty$ follows $\nu(x) = 0$.

[126.4.1] Suppose now that $a = 0$. *[126.4.2]* Then for every x and $\varepsilon > 0$

$$b - \varepsilon \leq a_n x + b_n \leq b + \varepsilon \tag{185}$$

[page 127, §0] if n is chosen sufficiently large. *[127.0.1]* By monotonicity of μ_n it follows that

$$\mu_n(b-\varepsilon) \le \mu_n(a_n x + b_n) \le \mu_n(b+\varepsilon). \tag{186}$$

[127.0.2] If ε is chosen so that $\mu(x)$ is continuous at the points $b-\varepsilon$ and $b+\varepsilon$, then

$$\mu(b-\varepsilon) \le \nu(x) \le \mu(b+\varepsilon). \tag{187}$$

[127.0.3] Because x was arbitrary it follows that $\mu(b-\varepsilon)=0$ and $\mu(b+\varepsilon)=1$. *[127.0.4]* Hence $\mu(x)$ is degenerate, contrary to the conditions above.

[127.1.1] Finally, let x be such that $\mu(x)$ is continuous at the point $ax+b$, and that $\nu(x)$ is continuous at x. *[127.1.2]* Then

$$\lim_{n\to\infty} \mu_n(a_n x + b_n) = \nu(x). \tag{188}$$

[127.1.3] On the other hand because $\lim_{n\to\infty}(a_n x + b_n) = ax+b$ one has for sufficiently large n that

$$ax+b-\varepsilon \le a_n x + b_n \le ax+b+\varepsilon, \tag{189}$$

where $\varepsilon > 0$ is chosen such that the distribution function μ is continuous at the points $ax+b-\varepsilon$ and $ax+b+\varepsilon$. *[127.1.4]* Hence by monotonicity

$$\mu_n(ax+b-\varepsilon) \le \mu_n(a_n x + b_n) \le \mu_n(ax+b+\varepsilon) \tag{190}$$

and for $n\to\infty$

$$\mu(ax+b-\varepsilon) \le \liminf_{n\to\infty} \mu_n(a_n x+b_n) \le \limsup_{n\to\infty} \mu_n(a_n x+b_n) \le \mu(ax+b+\varepsilon). \tag{191}$$

[127.1.5] Because $ax+b$ is a point of continuity for $\mu(x)$ and ε is arbitrary it follows that

$$\lim_{n\to\infty} \mu_n(a_n x + b_n) = \mu(ax+b) \tag{192}$$

and hence $\nu(x) = \mu(ax+b)$ proving the assertion.

References

[1] G.W. Scott-Blair and J.E. Caffyn. An application of the theory of quasi-properties to the treatment of anomalous stress-strain relations. *Phil. Mag.*, 40:80, 1949.

[2] H. Berens and U. Westphal. A Cauchy problem for a generalized wave equation. *Acta Sci. Math. (Szeged)*, 29:93, 1968.

[3] K.B. Oldham and J.S. Spanier. The replacement of Fick's law by a formulation involving semidifferentiation. *J. Electroanal. Chem. Interfacial Electrochem.*, 26:331, 1970.

[4] M. Caputo and F. Mainardi. Linear models of dissipation in anelastic solids. *Riv.Nuovo Cim.*, 1:161, 1971.

[5] K.B. Oldham and J.S. Spanier. *The Fractional Calculus*. Academic Press, New York, 1974.

[6] R. Hilfer. Classification theory for anequilibrium phase transitions. *Phys. Rev. E*, 48:2466, 1993.

[7] R. Hilfer. On a new class of phase transitions. In W.P. Beyermann, N.L. Huang-Liu, and D.E. MacLaughlin, editors, *Random Magnetism and High-Temperature Superconductivity*, page 85, Singapore,, 1994. World Scientific Publ. Co.

[8] R. Hilfer. Exact solutions for a class of fractal time random walks. *Fractals*, 3(1):211, 1995.

[9] R. Hilfer. Fractional dynamics, irreversibility and ergodicity breaking. *Chaos, Solitons & Fractals*, 5:1475, 1995.

[10] R. Hilfer. Foundations of fractional dynamics. *Fractals*, 3:549, 1995.

[11] R. Hilfer. An extension of the dynamical foundation for the statistical equilibrium concept. *Physica A*, 221:89, 1995.

[12] R.L. Bagley and P.J. Torvik. A theoretical basis for the application of fractional calculus to viscoelasticity. *J. Rheology*, 27:201, 1983.

[13] R.L. Bagley and P.J. Torvik. On the fractional calculus model of viscoelastic behaviour. *J. Rheology*, 30:133, 1986.

[14] W. Wyss. The fractional diffusion equation. *J. Math. Phys.*, 27:2782, 1986.

[15] W.R. Schneider and W. Wyss. Fractional diffusion and wave equations. *J. Math. Phys.*, 30:134, 1989.

[16] A.M.A. El-Sayed. Fractional-order diffusion-wave equation. *Int. J. Theor. Phys.*, 35:311, 1996.

[17] A. Compte. Stochastic foundations of fractional dynamics. *Phys.Rev. E*, 55:4191, 1996.

[18] T.F. Nonnenmacher and W.G. Glöckle. A fractional model for mechanical stress relaxation. *Phil. Mag. Lett.*, 64:89, 1991.

[19] L. Gaul, P. Klein, and S. Kempfle. Damping description involving fractional operators. *Mechanical Systems and Signal Processing*, 5:81, 1991.

[20] H. Beyer and S. Kempfle. Definition of physically consistent damping laws with fractional derivatives. *A. angew. Math. Mech.*, 75:623, 1995.

[21] R.R. Nigmatullin. The realization of the generalized transfer equation in a medium with fractal geometry. *phys. stat. sol. b*, 133:425, 1986.

[22] S. Westlund. Dead matter has memory ! *Physica Scripta*, 43:174, 1991.

[23] G. Jumarie. A Fokker-Planck equation of fractional order with respect to time. *J. Math. Phys.*, 33:3536, 1992.

[24] H. Schiessel and A. Blumen. Fractal aspects in polymer science. *Fractals*, 3:483, 1995.

[25] G. Zaslavsky. Fractional kinetic equation for Hamiltonian chaos. *Physica D*, 76:110, 1994.

[26] G. Zaslavsky. From Hamiltonian chaos to Maxwell's demon. *Chaos*, 5:653, 1995.

[27] E. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces*. Princeton University Press, Princeton, 1971.

[28] C. Berg and G. Forst. *Potential Theory on Locally Compact Abelian Groups*. Springer, Berlin, 1975.

[29] A.P. Prudnikov, Yu.A. Brychkov, and O.I. Marichev. *Integrals and Series*, volume 4. Gordon and Breach, New York, 1992.

[30] B.V. Gnedenko and A.N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, Cambridge, 1954.

[31] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume II. Wiley, New York, 1971.
[32] F. Oberhettinger. *Tables of Mellin Transforms.* Springer Verlag, Berlin, 1974.
[33] I.A. Ibragimov and Yu.V. Linnik. *Independent and Stationary Sequences of Random Variables.* Wolters-Nordhoff Publishing, Groningen, 1971.
[34] S. Bochner. *Harmonic Analysis and the Theory of Probability.* University of California Press, Berkeley, 1955.
[35] K. Yosida. *Functional Analysis.* Springer, Berlin, 1965.
[36] A.V. Balakrishnan. Fractional powers of closed operators and the semigroups generated by them. *Pacific J. Math.*, 10:419, 1960.
[37] U. Westphal. Ein Kalkül für gebrochene Potenzen infinitesimaler Erzeuger von Halbgruppen und Gruppen von Operatoren. *Compositio Math.*, 22:67, 1970.
[38] L. Bieberbach. *Lehrbuch der Funktionentheorie*, volume II. Teubner, Leipzig, 1931.
[39] W.R. Schneider. Completely monotone generalized Mittag-Leffler functions. *Expo. Math.*, 14:3, 1996.
[40] P.L. Butzer and R. J. Nessel. *Fourier Analysis and Approximation*, volume 1. Birkhäuser Verlag, Basel, 1971.
[41] R. Hilfer and L. Anton. Fractional master equations and fractal time random walks. *Phys.Rev.E, Rapid Commun.*, 51:848, 1995.
[42] R. Hilfer. On fractional diffusion and its relation with continuous time random walks. In A. Pekalski R. Kutner and K. Sznajd-Weron, editors, *Anomalous Diffusion: From Basis to Applications*, page 77. Springer, 1999.
[43] E.W. Montroll and G.H. Weiss. Random walks on lattices II. *J. Math. Phys.*, 6:167, 1965.
[44] M.N. Barber and B.W. Ninham. *Random and Restricted Walks.* Gordon and Breach Science Publ., New York, 1970.
[45] J.W. Haus and K. Kehr. Diffusion in regular and disordered lattices. *Phys.Rep.*, 150:263, 1987.
[46] J. Klafter, A. Blumen, and M.F. Shlesinger. Stochastic pathway to anomalous diffusion. *Phys. Rev. A*, 35:3081, 1987.
[47] B.D. Hughes. *Random Walks and Random Environments*, volume 1. Clarendon Press, Oxford, 1995.
[48] B.D. Hughes. *Random Walks and Random Environments*, volume 2. Clarendon Press, Oxford, 1996.
[49] J.K.E. Tunaley. Some properties of the asymptotic solutions of the Montroll-Weiss equation. *J. Stat. Phys.*, 12:1, 1975.
[50] E.W. Montroll and B.J. West. On an enriched collection of stochastic processes. In E.W. Montroll and J.L Lebowitz, editors, *Fluctuation Phenomena*, page 61, Amsterdam, 1979. North Holland Publ. Co.
[51] G.H. Weiss and R.J. Rubin. Random walks: Theory and selected applications. *Adv. Chem. Phys.*, 52:363, 1983.
[52] A. Erdelyi (et al.). *Higher Transcendental Functions*, volume III. Mc Graw Hill Book Co., New York, 1955.
[53] M.F. Shlesinger. Asymptotic solutions of continuous time random walks. *J. Stat. Phys.*, 10:421, 1974.
[54] M.F. Shlesinger, J. Klafter, and Y.M. Wong. Random walks with infinite spatial and temporal moments. *J. Stat. Phys.*, 27:499, 1982.
[55] C. Fox. The G and H functions as symmetrical Fourier kernels. *Trans. Am. Math. Soc.*, 98:395, 1961.
[56] B.L.J. Braaksma. Asymptotic expansions and anlytic continuations for a class of Barnes-integrals. *Compos.Math.*, 15:239, 1964.
[57] A.M. Mathai and R.K. Saxena. *The H-function with Applications in Statistics and Other Disciplines.* Wiley, New Delhi, 1978.
[58] H.M. Srivastava, K.C. Gupta, and S.P. Goyal. *The H-functions of One and Two Variables with Applications.* South Asian Publishers, New Delhi, 1982.
[59] A.P. Prudnikov, Yu.A. Brychkov, and O.I. Marichev. *Integrals and Series*, volume 3. Gordon and Breach, New York, 1990.
[60] I.N. Sneddon. *The Use of Integral Transforms.* Mc Graw Hill, New York, 1972.