Water Resources Research: 5, 1969, 260-267.

H14

Computer experiments with fractional Gaussian noises. Part 3: Mathematical appendix (M & Wallis 1969a)

• ***Chapter foreword.*** In effect, this appendix does little beyond restating M & Van Ness 1968 in a different style: less formal and more accessible to the original readers of Parts 1 and 2. Its original role was to allow a "big picture" to be drawn in the preceding two chapters without mathematical interruptions. M & Van Ness 1968 is now readily available as Chapter H11. But many readers will find that chapter difficult, making this text worth reprinting in slightly shortened form. Some long formulas were replaced by references to Chapter 11; self-similar by self-affine; and various awkward terms, by "bridge-range." •

BROWNIAN MOTION: DEFINITION, THE PROPERTY OF SELF-AFFINITY AND DIVERSE FORMS OF THE $\sqrt{\delta}$ LAW

First, let us recall some properties of the Brownian motion, a random process denoted by $B(t)$ and also called "Wiener Brownian motion," "Bachelier process" or "Wiener process." The main property of $B(t)$ is that, for every $\delta > 0$, the sequence of increments $B(t+\delta) - B(t)$ (defined when t is a multiple of δ) is a sequence of independent Gaussian random variables with zero mean and variance equal to δ. It is called a "discrete Gaussian white noise." Conversely, $B(t+\delta)$ (defined when t is a multiple of δ) is the cumulative sum of a sequence of independent Gaussian random variables. The function $B(t)$ is continuous (actually, "almost surely continuous," but we may disregard this technicality). Therefore, $B(t)$ can be considered a tool for starting with a function of discrete time (namely, the above mentioned cumulative sum of independent Gaussian

random variables) and then interpolating it into a function of continuous time.

The Wiener Brownian motion is a "self-affine process" and satisfies "$\sqrt{\delta}$ laws." To explain these terms, change the unit of time in the positive but otherwise arbitrary ratio r. The new time u and the old time t are related by $t = ru$. Then, a fundamental property of Wiener's Brownian motion is as follows:

"The rescaled function $r^{-0.5}[B(ru) - B(0)]$ is identical in distribution to the original function $B(t) - B(0)$".

First corollary of self-affinity. The variance of $[B(t+\delta) - B(t)]/\sqrt{\delta}$ is defined as equal to 1. Thus, we obtain the result that

"Standard deviation of $B(t+s) - B(t) = \sqrt{\delta}$ for every t and δ."

This property is called a "uniform $\sqrt{\delta}$ law for the standard deviation." The term "uniform" stands for "valid for every t and for every δ."

Second corollary of self-affinity. It concerns the range

$$R_P(t, \delta) = \max_{0 \le u \le \delta} [B(t+u) - B(t)] - \min_{0 \le u < \delta} [B(t+u) - B(t)],$$

with a *continuously* varying "dummy variable" u. We have

"The distribution of the variable $R_P(t, \delta)/\sqrt{\delta}$ is independent of t and δ."

This property is called a "uniform $\sqrt{\delta}$ law for the population range." The moments of the random variable $R_P(t, \delta)$ have been evaluated in Feller 1951.

Third corollary of self-affinity. It concerns the bridge range

$$R(t, \delta) = \max_{0 \le u \le \delta} \{[B(t+u) - B(t)] - (u/\delta)[B(t+\delta) - B(t)]\}$$
$$- \min_{0 \le u \le \delta} \{[B(t+u) - B(t)] - (u/\delta)[B(t+\delta) - B(t)]\},$$

with a continuously varying dimension variable u. We have

"The distribution of the variable $R(t, \delta)/\sqrt{\delta}$ is independent of t and δ."

The moments of $R(t, \delta)$ have also been evaluated in Feller 1951: where as expected, $\mathscr{E}[R(t, \delta)/\sqrt{\delta}\,] < \mathscr{E}[R_p(t, \delta)/\sqrt{\delta}\,]$ and the scatter around the expectation is smaller for $R(t, \delta)$ than for $R_p(t, \delta)$.

FRACTIONAL BROWNIAN MOTION: DEFINITIONS, SELF-AFFINITY, AND THE δ^H LAWS

From the mathematical viewpoint, fractional Brownian motions with $H \neq 0.5$ are characterized by being Gaussian, of stationary increments and self-affine:

(a) they are *Gaussian processes;*

(b) their increments $B_H(t) - B_H(t - \delta)$, defined when t is a multiple of δ, constitute a *stationary* random process; and

(c) they are *self-affine,* in the sense that

"If time is changed in the ratio r, then the rescaled function $r^{-H}[B_H(tr) - B_H(0)]$ is identical in distribution to the original $B_H(t) - B_H(0)$."

It can be shown that such a function $B_H(t) - B_H(0)$ exists if and only if $0 < H < 1$ and that $B_H(t)$ can be obtained from Wiener's Brownian $B(t)$ by forming

$$B_H(t) - B_H(0) = \int_{-\infty}^{0} [(t-u)^{H-0.5} - (-u)^{H-0.5}]dB(u) + \int_{0}^{t} (t-u)^{H-0.5}dB(u).$$

This integral transforms all the values of the process $B'(t)$, corresponding to all instants u prior to t, by weighting and then adding them. The result can be shown to be a continuous function of t (actually, "almost surely continuous," but – again – we may disregard this technicality).

The above representation of $B_H(t)$ may appear artificial. To make it symmetric, it is tempting to write

$$B_H(t') - B_H(t'') = \int_{-\infty}^{t'} (t'-u)^{H-0.5}dB(u) - \int_{-\infty}^{t''} (t''-u)^{H-0.5}dB(u).$$

Unfortunately, both integrals on the right-hand side are divergent, so that this formula is useful mainly as a mnemonic device: fractional Brownian motion is formed by weighting past values of a white noise by $(t-u)^{H-0.5}$.

It is clear that fractional Brownian motion with $H = 0.5$ reduces to

$$B_{0.5}(t) - B_{0.5}(0) = \int_0^t dB(u) = B(t) - B(0).$$

Thus, $B_{0.5}(t)$ is ordinary Brownian motion, and fractional Brownian motion generalizes ordinary Brownian motion to values of H different from $H = 0.5$.

First corollary of the self-affinity of fractional Brownian motion. It follows from the fact that the variance of $\delta^{-H}[B_H(t+\delta) - B_H(t)]$ is independent of t and δ. Thus,

$$\mathscr{E}[B_H(t+\delta) - B_H(t)]^2 = \mathscr{E}[\Delta B_H]^2 = C_H \delta^{2H}$$

for every t and δ. The constant C_H is written down in Chapter 11. In other terms:

"Standard deviation of $B_H(t+\delta) - B_H(t) = \sqrt{C_H}\,\delta^H$ for every t and δ."

Of these two equivalent statements, the first was used in the discussion of "terms" in M & Wallis 1969a{H12}. The second is a "uniform δ^H law for the standard deviation," to be used in discussing the corollary to Hurst's law due to Langbein 1956.

Second corollary of self-affinity. It concerns the range

$$R_P(t, \delta) = \max_{0 \le u \le \delta} [B_H(t+u) - B_H(t)] - \min_{0 \le u \le \delta} [B_H(t+u) - B_H(t)].$$

"The distribution of the variable $R_P(t, \delta)\delta^{-H}$ is independent of t and δ."

This is a "uniform δ^H law for the population range." In particular, D_H being a constant different from C_H, one has

$$\mathscr{E}[R_P(t, \delta)] = D_H \delta^H \text{ for every } t \text{ and } \delta.$$

Third corollary of self-affinity. It concerns the corresponding bridge range $R(t, \delta)$.

"The distribution of the variable $R(t, \delta)\delta^{-H}$ is independent of t and δ."

It appears most plausible that, generalizing Feller's result valid for $H = 0.5$, one has $\mathcal{E}[R(t, \delta)\delta^{-H}] < \mathcal{E}[R_p(t, \delta)\delta^{-H}]$, and that the scatter around the expectation is smaller for $R(t, \delta)$ than for $R_p(t, \delta)$.

DISCRETE-TIME FRACTIONAL NOISES: DEFINITIONS

By analogy with white noise, the sequence of increments of $B_H(t)$, namely the sequence of values of $\Delta B_H(t) = B_H(t) - B_H(t-1)$, with t an integer, is called "discrete-time fractional Gaussian noise." It can be deduced from a Brownian motion $B(u)$ by the formula

$$\Delta B_H(t) = B_H(t) - B_H(t-1) = \int_{-\infty}^{t} K_H(t-u)dB(u),$$

with the "kernel function" $K_H(u)$ given by

$$K_H(u) = u^{H-0.5} \qquad \text{for } 0 < u < 1,$$
$$K_H(u) = [u^{H-0.5} - (u-1)^{H-0.5}] \qquad \text{for } u > 1.$$

The covariance of $\Delta B_H(t)$ takes the following form:

$$\text{if } \delta = 0, \quad \mathcal{E}[\Delta B_H(t)\Delta B_H(t+\delta)] = \mathcal{E}[\Delta B_H^2] = C_H\delta^{2H},$$

$$\text{if } \delta > 0, \quad \mathcal{E}[\Delta B_H(t)\Delta B_H(t+\delta)] = \mathcal{E}[\Delta B_H(1)\Delta B_H(1+\delta)]$$

$$= (C_H/2)[(\delta+1)^{2H} - 2\delta^{2H} + (\delta-1)^{2H}].$$

Clearly, if $X(t) = \Delta B_H(t) = B_H(t) - B_H(t-\delta)$, one has

$$X_\Sigma(t) = \sum_{u=1}^{t} [B_H(u) - B_H(u-1)] = B_H(t) - B_H(0).$$

"The standard deviation of $X_\Sigma(t+\delta) - X_\Sigma(t) = \sqrt{C_H}\,\delta^H$ for all *integer* values of t and δ."

This is the form of the "δ^H law for the standard deviation." It is not valid uniformly but only for integer t and δ.

In the section following the next, we shall encounter additional complications in identifying the δ^H law applicable to the range of discrete-time fractional noise.

WHITE AND FRACTIONAL GAUSSIAN NOISES IN CONTINUOUS TIME

The derivatives $B'(t)$ or $B'_H(t)$ are called, respectively, "continuous-time white noise" and "continuous-time fractional noise." Extraordinary care is required in handling these expressions. The reason is that, although $B_H(t)$ is a continuous function of t, its local behavior is extremely irregular. It may be recalled from calculus that when a function $X(t)$ is differentiable at the point t, its derivative is defined as the limit for $\delta \to 0$ of the finite local slope. That is,

$$X'(t) = \lim_{\delta \to 0} \delta^{-1}[X(t+\delta) - X(t)].$$

However, the local behavior of fractional Brownian motion is so irregular that $\delta^{-1}[B_H(t+\delta) - B_H(t)]$ *does not tend towards any limit* as $\delta \to 0$. Thus, $B_H(t)$ is continuous but has no derivative in the ordinary sense.

Perhaps surprisingly, this seemingly "pathological" feature is needed to make $B_H(t)$ a realistic model of geophysical phenomena. For example, total precipitation between times 0 and t must be a continuous function t, and such is indeed the case for $B_H(t)$. But the "instantaneous" precipitation is extraordinarily variable in its detail. One would not even want to follow it in continuous time, except after it has been appropriately smoothed. To model instantaneous precipitation by the formal derivative $B'_H(t)$ is acceptable precisely *because* this $B'_H(t)$ becomes mathematically meaningful only after it is agreed that it must always be examined after some smoothing mechanism has eliminated the high-frequency jitter.

COMPARISON OF THE RANGES FOR DISCRETE AND CONTINUOUS-TIME FRACTIONAL NOISES: GRID CORRECTION

By the definitions of the last two sections,

$$B_H(t) - B_H(0) = \int_0^t B'_H(u)du = \sum_{u=1}^{t} [B_H(u) - B_H(u-1)].$$

Thus, for both discrete and continuous time, the range of a fractional noise is

$$R_P(t, \delta) = \max_{0 \le u \le \delta} \big[B_H(t+u) - B_H(t)\big] - \min_{0 \le u \le \delta} \big[B_H(t+u) - B_H(t)\big].$$

The difference between discrete and continuous time concerns the acceptable values of u. In discrete time, one deals with expressions obtained by comparing $\delta + 1$ quantities of the form $B_H(t+u) - B_H(t)$. The relevant quantities are the "grid max," defined as $\max_{0 \le u \le \delta}[B_H(t+u) - B_H(t)]$, the corresponding "grid min," and their difference, to be called the "grid R_P." But in continuous time, one deals with expressions obtained by comparing an infinite number of quantities, among which the above $\delta + 1$ quantities are included. The resulting expressions will be called "true max," "true min" and "true R_P." Consider the process that selects max, min or R_P as a "comparison shopper." When time is made continuous, the selection open to our shopper becomes wider so that "true max" will be at least as large (and possibly larger) than the "grid max" that would have been previously selected. This argument, and its obvious counterparts for min and R_P, yields the following inequalities

$$\begin{aligned} \text{"true max"} - \text{"grid max"} &\ge 0 \\ \text{"true min"} - \text{"grid min"} &\le 0 \\ \text{"true } R_P\text{"} - \text{"grid } R_P\text{"} &\ge 0. \end{aligned}$$

To make these inequalities even more persuasive, let us restate the first differently. Denote by $t + u_{\max}$ the precise instant in discrete time where $B_H(t+u) - B_H(t)$ attains its "grid max." Assuming, in addition, that $\delta > 2$ and $1 \le u_{\max} \le \delta - 1$, this definition implies that

$$B_H(t + u_{\max}) \ge B_H(t + u_{\max} - 1) \text{ and } B_H(t + u_{\max}) \ge B_H(t + u_{\max} + 1).$$

Moreover, it would be extraordinarily unlikely for $B_H(t+u) - B_H(t)$ to have a sharp continuous time peak at the precise instant $t + u_{\max}$. It follows that there is a point in the interval from $t + u_{\max} - 1$ to $t + u_{\max} + 1$ where $B_H(t+u) - B_H(t)$ attains a local maximum higher than $B_H(t + u_{\max}) - B_H(t)$. Since "true max" is at least equal to that local maximum, we have proved the inequality "true max" – "grid max" ≥ 0.

In fact, one can prove much more. Except when δ is very small, say for $\delta \ge 3$, the distributions of the differences "true max" – "grid max" and "true R_P" – "grid R_P" will depend very little upon δ. Since R_P tends to

increase with δ, the relative error due to the grid, defined as "true R_p"/"grid R_p" -1, will tend to decrease as $\delta \to \infty$. For large δ, this relative grid error will be small.

Moreover, as δ increases from 1 to ∞, $\mathscr{E}[R_p(t, \delta)\delta^{-H}]$ will increase from 0 to the value D_H that enters into the δ^H law for the range of $B_H(t)$ in continuous time. Thus, the graph of $\log \mathscr{E}[R_p(t, \delta)]$ versus $\log \delta$ will be a straight line of slope H, preceded by a portion with strong downward concavity. We have verified numerically that such is indeed the case.

Unfortunately, the precise position of the bend between this initial transient and the asymptotic δ^H behavior is difficult to determine by analysis. If this bend was positioned at an extremely large value of δ, the δ^H asymptote would have no relevance to modeling. Our computer experiments demonstrate that, even for simple approximations to fractional noise, this bend is positioned at a smaller value of δ.

The need for a grid correction for small δ and the resulting complicated transient result in a more complicated theory. The size of the grid correction could readily be decreased by selecting a finer grid for u. But we are attempting to model physical reality, and we know that records are necessarily taken in discrete time. To be comparable with the data, a theory must use a discrete grid; units of time can then be chosen so that the theoretical and practical grids give identical results.

THE BRIDGE RANGE AND THE RESCALED BRIDGE RANGE

The bridge range $R(t, \delta)$ is more complicated than the range $R_p(t, \delta)$ and will not be discussed here. The ratio $R(t, \delta)/S(t, \delta)$ is even more complicated. Due to partial compensation between the grid correction and the correction due to division by $S(t, \delta)$, the ratio $R(t, \delta)/S(t, \delta)$ has a smaller initial transient than R_p.

DEFINITION OF A TYPE 1 APPROXIMATE FRACTIONAL NOISE

The numerical evaluation of the integral $\int_{-\infty}^{t} K_H(t-u)dB(u)$ on a computer involves diverse approximations:

1. The infinite span $-\infty < u < t$ must be replaced by a finite span of the form $-M + t < u < t$. The finite M introduces a *low frequency* error term. Its most striking consequence is to make the asymptotic behavior of $R(t, \delta)/S(t, \delta)$ scale like $\sqrt{\delta}$ rather than δ^H.

2. A discrete increment ε must be selected for the variable of integration u; in particular, the infinitesimal $dB(u)$ must be replaced by a finite difference $B(u+\varepsilon) - B(u)$. The positive ε introduces a *high frequency* error term.

3. Once the increment $\varepsilon > 0$ is selected, there is little point in computing the kernel K_H exactly. Simplification of the analytic form of K_H need not introduce additional major errors.

4. Finally, another discrete increment must be selected to compute $K(t-u)[B(u+\varepsilon) - B(u)]$ to a finite number of decimals.

Two typical approximations to $B_H(t)$, satisfying the above conditions, will be described in this and the next sections.

Our first and foremost approximation to fractional noise leads to Type 1 functions. Steps (1) and (2) are taken by selecting $\varepsilon = 0.10$. Steps (3) and (4) are then satisfied as follows. The continually varying kernel $K_H(u)$ is replaced by the stepwise varying kernel $K_1(u \mid H, M)$ that is illustrated in Figure 1 and is described as follows:

- When $u > M$, $K_1(u \mid H, M) = 0$.
- When $u < M$ and u is an exact multiple of $1/10$, $K_1(u \mid H, M) = K_H(u)$.
- When $u < M$ and u is not an exact multiple of $1/10$ and when w is the smallest multiple of $1/10$ whose value exceeds u, $K_1(u \mid H, M) = K_H(w)$.

This kernel is used to define the Type 1 approximation

$$F_1(t \mid H, M) = \int_{-0}^{M} K_1(t-u \mid H, M) dB(u).$$

To consider the high-frequency aspects of the function F_1, set $M = \infty$. The kernel $K_1(u \mid, H, \infty)$ is a good approximation to $K_H(u)$ when H is such that $\Delta B_H(t)$ does not have very strong high-frequency components. Such is the case when $K_H(u)$ is continuous, that is, if $0 < H < 0.5$. Generally, the approximation of $K_H(u)$ by $K_1(u \mid H, \infty)$ improves as H increases from 0 to 1. This fits the results of our experimental study of the behavior of $R(t, \delta)/S(t, \delta)$ for Type 1 functions.

Of course, there is nothing exclusive about the value 0.10 for ε. A smaller ε improves the approximation of $K_H(u)$ by $K_1(u \mid H, \infty)$ for all values of H and particularly for small H. In other words, to achieve the same quality of approximation for every value of H, it is necessary to make ε a function of H, with $\varepsilon(H)$ decreasing as H increases from 0 to 1.

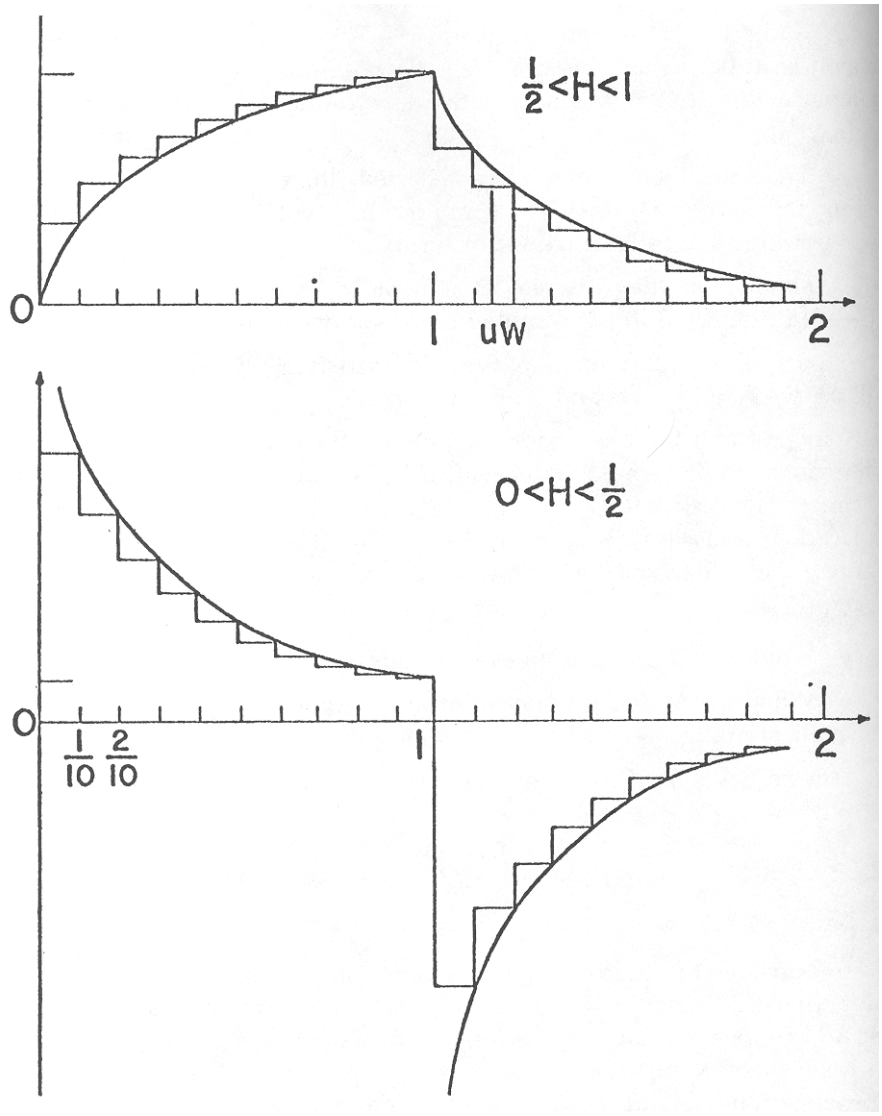

FIGURE C14-1. Relation between the kernels K_H and K_1. On both portions of the figure, the bold lines represent the overall shapes of the kernel $K_H(u)$. The thin lines illustrate the definition of w and of the stepwise, varying, approximating kernel $K_1(u|H, M)$. "Staircase-shaped" curves were drawn, but the "risers" were inserted solely for legibility; only the "treads" count. These step functions correspond to representations of the relevant random functions by moving average *integrals* in continuous time. To obtain the representations by moving average *sums*, one should replace each tread by its rightmost point.

The additional low-frequency error introduced by setting $M < \infty$ is essentially independent of ε.

A moving average representation. While our kernel K_1 is stepwise varying, the function F_1 is evaluated only when t is an integer. For such values of t, our Type 1 function is a discrete-time moving average of the form

$$F_1(t \mid H, M) = \sum_{n=10(t-M)}^{10t-1} K_1(t - n/10 \mid H, M) G(n/10).$$

In this formula, the quantities $G(n/10)$, defined for all integer n, are independent Gaussian variables with zero mean and variance $1/10$. The kernel K_1 has the same analytic form as defined previously; but, of course, the argument $n/10$ in the kernel is now a multiple of $1/10$ so that the complications relative to w no longer arise.

It may be observed that in moving average integrals both the time t and the dummy variable u vary continuously. In the usual moving average sums, the time t and the dummy variable u are both integers. On the contrary, in the above definition of a Type 1 function, the time t is an integer, but the dummy variable u is a multiple of 0.1, which places it somewhere between an integer and a continuously variable quantity.

DEFINITION OF A TYPE 2 APPROXIMATE FRACTIONAL NOISE

Type 1 functions require quite extensive computations. To simplify them, one could select a larger ε, while also replacing $K_H(u)$ by some more manageable function, as suggested by step (3) in the preceding section's approximations. For large u, $K_1(u \mid H, \infty) \sim (H - 0.5)u^{H-1.5}$. This function becomes infinite for $u = 0$ so that a different analytic formula must be selected near $u = 0$. Different approaches must be followed when $0 < H < 0.5$ and $0.5 < H < 1$.

The case $0.5 < H < 1$. We shall define a Type 2 approximating kernel $K_2(u \mid H, M)$ as follows:

- If $u > M$, $K_2(u \mid H, M) = 0$.
- If $u < M$ and u is an integer, $K_2(u \mid H, M) = (H - 0.5)u^{h-1.5}$.
- If $u < M$ and u is not an integer and if w is the smaller integer whose value exceeds u,

$$K_2(u \mid H, M) \sim (H - 0.5)w^{H-1.5}.$$

The discrete-time moving average corresponding to this kernel is

$$F_2(t \mid H, M) = (H - 0.5) \sum_{u=t-M}^{t-1} (t - u)^{H-1.5} G(u),$$

which is the Type 2 approximation already described at the very beginning of this paper.

Type 1 and Type 2 functions are essentially identical at very low-frequencies, but have different high-frequencies. Therefore, Type 1 approximations may be expected to be best when high-frequency effects are slight, that is, when H is near 1. We saw in our discussion of R/S that the two approximations indeed differ little for $H = 0.9$. However, when $H = 0.5$, the kernel K_1 vanishes except for $u < 1$, whereas K_2 remains nonzero throughout $u < M$. Therefore, the two approximations differ greatly for $H = 0.5$, and K_2 is of limited scope. Its usefulness is due to formal simplicity, combined with the fact that empirical values of H are often near 1.

The case $0 < H < 0.5$. Here, the high-frequency effects are very strong, as shown in Figure 17 of M & Wallis 1969b{H27}. For this reason, Type 2 approximations are a bit more complicated to implement. In writing the approximate kernel K_2, one must ensure that $\sum_{u=0}^{\infty} K_2(u \mid H, M)$ either vanishes or is very close to zero. This is the purpose behind the positive factor $Q_H G(0)$ introduced in the formula at the beginning of Chapter 12.