

SOME OBSERVATIONS ON ALGORITHMS OF THE GAUSS–BORCHARDT TYPE

by JAAK PEETRE

(Received 3rd December 1989)

A one parameter family of algorithms is studied, which contains both the arithmetic-geometric mean of Gauss and its generalization by Borchardt, recently studied by J. and P. Borwein. We prove that the presence of an asymptotic formula for such an algorithm is, in view of the Poisson summation formula, equivalent to the vanishing of certain integrals. In the case of Gauss and Borchardt the latter involve theta functions. Finally, we investigate the question of convergence of the algorithm for complex values, thereby generalizing the corresponding result of Gauss.

1980 *Mathematics subject classification* (1985 Revision): 26A18, 33A25.

Introduction

As is well-known, the arithmetic-geometric mean, studied by Gauss [8] and others, arises when iterating the map

$$a, b \mapsto a' = \frac{a+b}{2}, b' = \sqrt{ab}. \quad (1)$$

As a generalization Borchardt [2] considered iterations of the 4-dimensional map

$$a, b, c, e \mapsto a' = \frac{a+b+c+e}{4}, b' = \frac{\sqrt{ab} + \sqrt{ce}}{2}, c' = \frac{\sqrt{ac} + \sqrt{be}}{2}, e' = \frac{\sqrt{ae} + \sqrt{bc}}{2}. \quad (2)$$

When putting $b=c=e$ (and $b'=c'=e'$) in (2) there arises from (2) the 2-dimensional map

$$a, b \mapsto a' = \frac{a+3b}{4}, b' = \frac{\sqrt{ab} + b}{2}. \quad (2')$$

The resulting mean has recently been thoroughly investigated in a beautiful paper by J. and P. Borwein [4]. (We understand that the same authors intend to treat more general “higher order” algorithms in a forthcoming monograph [6].) They show that it has properties entirely parallel to the ones of the arithmetic-geometric mean. In particular,

in both cases one has an asymptotic formula, viz. denoting the mean (that is, the limit of the iteration) by M holds

$$M(a, b) \sim \frac{\pi}{2} \cdot \frac{b}{\log(a/b)}, \quad b \rightarrow 0 \quad (3)$$

in case (1) and

$$M(a, b) \sim \frac{\pi^2}{3} \cdot \frac{b}{(\log(a/b))^2}, \quad b \rightarrow 0 \quad (4)$$

in case (2'). Now both maps can be imbedded in a one parameter family of maps

$$a, b \mapsto a' = \frac{a + (\lambda - 1)b}{\lambda}, \quad b' = \frac{2\sqrt{ab} + (\lambda - 2)b}{\lambda}, \quad (5)$$

taking $\lambda = 2$ we get back (1) and taking $\lambda = 4$ we get back (2'). It seems—but up to our knowledge this has not been rigorously established yet—that the expected analogue (generalization) of (3) and (4) fails except in these two special cases.

The map (5) has been discussed previously by us in [1, 10, 11]. In the present note we study a general class of iterations mimicking in various ways the behaviour of the arithmetic-geometric mean (1) and its generalization (2'), (4). In particular, we show that in view of Poisson's summation formula the presence of an asymptotic formula of type (3), (4) is equivalent to vanishing of certain integrals depending on a discrete parameter. Moreover, we establish a kind of "duality". It turns out that the maps of the type (5) are self-dual in this sense. This is well-known in the theory of arithmetic-geometric mean [8] and is also inherent in the treatment in [4].

All this is laid out in Sections 1–4.

In Section 5, which is by and large independent of the rest of the paper, we study the convergence of the "generalized arithmetic-geometric mean" (5) for complex values of a, b . We prove there a partial extension of Gauss's classical results about the existence of the limit. In the case of the "ordinary" arithmetic-geometric mean (1) Gauss was able to connect the various "determinations" of the arithmetic-geometric mean through the theory of the modular function (see [4] for a contemporary treatment, and also some of the references given there). Of course, it is hardly probable that something similar is possible in general, except of course in the case $\lambda = 4$ treated by Borchardt [2] and the Borweins [3, 4].

Acknowledgement. The author is indebted to the referee for pointing out a number of inaccuracies in the original version of the paper.

1. A class of algorithms

A very general class of algorithms is obtained by iterating maps of the type

$$a, b \mapsto a' = f(a, b), b' = g(a, b), \tag{1}$$

where f and g are homogeneous functions of degree 1 (see e.g. [1]). It is then natural to use, instead of the variables a, b , assumed to be positive, the inhomogeneous parameter $x = b/a$. We will concentrate on the case $b < a$ so $0 < x < 1$. Then one gets the one dimensional map

$$x \mapsto x' = \phi(x) \stackrel{\text{def}}{=} \frac{g(x)}{f(x)} \tag{2}$$

where we have written $g(x) = g(1, x)$, $f(x) = f(1, x)$. The corresponding mean M is invariant under the map in (2), which in inhomogeneous notation means that

$$M(x) = f(x)M(\phi(x)), \tag{3}$$

where again $M(x) = M(1, x)$.

The inhomogeneous data ϕ, f completely determine the map in (1).

Example 1.

$$\phi(x) = \frac{2\sqrt{x} + (\lambda - 2)x}{1 + (\lambda - 1)x}, f(x) = \frac{1 + (\lambda - 1)x}{\lambda}$$

in the case of the generalized arithmetic-geometric mean, that is, the mean corresponding to the map (5) in the Introduction. In particular,

$$\phi(x) = \frac{2\sqrt{x}}{1+x}, f(x) = \frac{1+x}{2} \text{ if } \lambda = 2 \text{ (Gauss's case)}$$

and

$$\phi(x) = \frac{2\sqrt{x} + 2x}{1 + 3x}, f(x) = \frac{1 + 3x}{4} \text{ if } \lambda = 4 \text{ (Borchardt's case).}$$

Now we restrict the situation further by imposing some conditions on ϕ, f . As we are only aiming at being able to incorporate certain very concrete examples (mainly the algorithm in formula (5) in the Introduction, to which we will henceforth informally refer to as the *generalized arithmetic-geometric mean*), we need not be very explicit about the assumptions made. Roughly speaking, our policy is to assume that all functions are “analytic”, meaning that they admit, when necessary, suitable expansions in powers of the variables. First of all we assume that $\phi: (0, 1) \rightarrow (0, 1)$ is a diffeomorphism (or perhaps, better, a “bianalytic” such). Furthermore 1 is supposed to be a (quadratically) *hyperattractive* fixed point, meaning that $\phi(x) = 1 + c_1(x - 1)^2 + \dots$, and, similarly, 0 is

supposed to be a (quadratically) *hyper-repelling* fixed point, meaning that $\phi^{(-1)}(x) = c_0x^2 + \dots$. (Here and in the sequel $\phi^{(-1)}$ stands for the inverse function. A similar notation is used for iterates.) About the function f we shall assume that $f: [0, 1] \rightarrow [\lambda^{-1}, 1]$ is a diffeomorphism for some $\lambda > 1$, so that we have the expansions $f(x) = 1 + d_1(x-1) + \dots$ and

$$f(x) = \frac{1}{\lambda} + d_0x + \dots$$

Note that $d_0, d_1 > 0$.

Remark 1. In the assumptions made clearly, the points 0 and 1 play to some extent a symmetric rôle. We shall therefore try to adhere to the convention, as was just made, that a subscript 0 refers to the point 0 and, similarly, that a subscript 1 refers to the point 1.

Note thus that the following considerations are somewhat formal, but it is clear that they are applicable in the examples of interest to us, this being our main ambition.

Proposition 1. *The functional equation (3) has a unique solution M continuous at $x = 1$ such that $M(1) = 1$.*

Proof. It is clear that the infinite product

$$M(x) \stackrel{\text{def}}{=} \prod_{n=0}^{\infty} f(\phi^{(n)}(x)) \tag{4}$$

gives such a solution. Conversely, assume that there are two solutions M_1 and M_2 . Then their quotient $M = M_1/M_2$ is ϕ -invariant, $M(x) = M(\phi(x))$, and continuous at $x = 1$ with $M(1) = 1$. But then M is constant, $M \equiv 1$, and $M_1 = M_2$. \square

We wish to investigate the behaviour of this canonical solution M as $x \rightarrow 0$. To this end we consider the functional equation

$$Z(x) = \lambda f(x)Z(\phi(x)). \tag{5}$$

We then have:

Proposition 2. *The functional equation (4) has a unique solution Z continuous at $x = 1$ such that $Z(1) = 1$.*

Proof. Same proof as for Proposition 1 but now the solution is equivalently given explicitly by the formula

$$Z(x) \stackrel{\text{def}}{=} \prod_{n=-\infty}^{-1} (\lambda f(\phi^{(n)}(x)))^{-1}. \quad \square \tag{6}$$

Note that ((4)+(6)!)

$$\frac{M(x)}{Z(x)} = \prod_{n=-\infty}^{\infty} \lambda_n f(\phi^{(n)}(x)), \tag{7}$$

with $\lambda_n = 1 (n \geq 0)$, $\lambda_n = \lambda (n < 0)$.

Next we invoke the functional equation

$$(N(\phi(x)))^2 = N(x), \tag{8}$$

which is essentially *Böttcher’s equation* for the map $\phi^{(-1)}$; to see this we have to rewrite (8) as

$$(N(x))^2 = N(\phi^{(-1)}(x)). \tag{8’}$$

The following is well-known.

Proposition 3. *If 0 is a (quadratically) attractive fixed point for the map $\phi^{(-1)}$ then equation (8) or (8’) has a unique solution such that $N(x) \approx c_0 x$ as $x \rightarrow 0$.*

Proof. Indeed, we have the explicit formula

$$N(x) = \lim_{n \rightarrow \infty} (c_0 \phi^{(-n)}(x))^{(1/2^n)}. \tag{9}$$

(This formula again formally arises by writing (8’) as

$$N(x) = (N(\phi^{(-1)}(x)))^{1/2} = (N(\phi^{(-2)}(x)))^{1/4} = \dots,$$

along with the approximation $N(x) \approx c_0 x$.) \square

Remark 2. Alternatively, writing $\phi^{(-1)}(x) = c_0 x^2 P(x)$ we may rewrite (9) as (cf. [4])

$$N(x) = c_0 x \cdot P(x)^{1/2} \cdot P(\phi^{(-1)}(x))^{1/4} \cdot \dots$$

We remark that the range of N , apparently, is $(0, 1)$. The intuitive meaning of the function N is that we can use $q = N(x)$ as a “uniformizing” parameter near 0: an application of $\phi^{(-1)}(x)$ affects the change $q \mapsto q^2$, while an application of ϕ itself causes $q \mapsto q^{1/2}$.

Returning to the original function equation (3) we see, in view of (5), that (3) has, besides the function M given by the infinite product (4) in the proof of Proposition 1, a second solution M' defined by

$$M'(x) = \frac{Z(x)}{(\log(1/N(x)))^{(\log \lambda / \log 2)}}.$$

Indeed, we get, using (5) and (8),

$$f(x)M'(\phi(x)) = \frac{f(x)Z(\phi(x))}{(\log(1/N(\phi(x))))^{(\log \lambda / \log 2)}}$$

$$\begin{aligned}
 &= \frac{\lambda^{-1} Z(x)}{(\log((1/N(x)))^{1/2})^{(\log \lambda/\log 2)}} = \frac{\lambda^{-1} Z(x)}{(1/2)^{(\log \lambda/\log 2)} (\log(1/N(x)))^{(\log \lambda/\log 2)}} \\
 &= \frac{Z(x)}{(\log(1/N(x)))^{(\log \lambda/\log 2)}} = M'(x). \quad \square
 \end{aligned}$$

In particular, the quotient $(M(x)/M'(x))$ is ϕ -invariant.
 When is this quotient a constant? If this is the case, we can write

$$M(x) = A \frac{Z(x)}{(\log(1/N(x)))^{(\log \lambda/\log 2)}}.$$

In particular, we have then an asymptotic formula

$$M(x) \sim A \frac{1}{(\log(1/c_0 x))^{(\log \lambda/\log 2)}}, \quad x \rightarrow 0$$

(cf. (3) and (4) in the Introduction). In the next Section we will give a necessary and sufficient condition, in the form of the vanishing of infinitely many integrals, for this to be true.

2. Application of Poisson’s summation formula

This formula states that if F is any function on the real line $\mathbf{R} = (-\infty, \infty)$ then, with appropriate assumptions and with appropriate interpretation of the summation, holds

$$\sum_{n \in \mathbf{Z}} F(t + 2\pi n) = \frac{1}{2\pi} \sum_{n \in \mathbf{Z}} \hat{F}(n) e^{itn},$$

i.e. the left hand side is a periodic function of period 2π whose Fourier coefficients are obtained by restricting the Fourier transform \hat{F} of F to the integers \mathbf{Z} . Here, formally, the Fourier transform of F is defined as

$$\hat{F}(\xi) = \int_{\mathbf{R}} e^{-i\xi t} F(t) dt,$$

with the inverse

$$F(t) = \frac{1}{2\pi} \int_{\mathbf{R}} e^{i\xi t} \hat{F}(\xi) d\xi, \quad (\text{Fourier’s inversion formula}).$$

In particular, the left hand side is independent of t iff $\hat{F}(n) = 0$ ($n \in \mathbf{N}$), provided that F is real valued.

In our case, the natural parameter to be taken is

$$t = \frac{2\pi \log \log(1/N(x))}{\log 2}.$$

(Upon applying the map ϕ (respectively $\phi^{(-1)}$), t experiences a change by -2π (respectively 2π) units. If x ranges from 0 to 1 then $\log N(x)$ ranges from ∞ to 0, while t ranges from ∞ to $-\infty$.)

We assume that we are in the situation that there is an asymptotic formula, i.e.

$$M(x) \sim A \frac{1}{(\log(1/c_0 x))^{\log \lambda / \log 2}}, \quad x \rightarrow 0 \quad (A = \text{constant}). \tag{1}$$

Then by taking logarithms we obtain from (9) in Section 1

$$\sum_{n \in \mathbb{Z}} \log f(\phi^{(n)}(x)) + \log \lambda_n + \log \lambda \cdot \frac{\log \log(1/N(x))}{\log 2} = \text{constant}$$

or, differentiating with respect to t ,

$$\sum_{n \in \mathbb{Z}} \frac{d}{dt} \log f(\phi^{(n)}(x)) + \frac{1}{2\pi} \log \lambda = 0.$$

If Poisson’s summation formula is applicable and all other formalities are justified, we therefore may conclude (see the preceding discussion) that

$$\int_{\mathbb{R}} \frac{d}{dt} \log f(x) e^{int} dt = 0 \quad (n \in \mathbb{N})$$

or, by change of variables,

$$\int_0^1 \left(\log \frac{1}{N(x)} \right)^{2\pi in / \log 2} d \log f(x) = 0 \quad (n \in \mathbb{N}). \tag{2}$$

We are thus able to conclude:

Theorem 1. *The validity of the asymptotic formula is equivalent to the vanishing of the infinitely many integrals in (2).*

In the case of Gauss and Borchardt all the relevant data are expressible in terms of theta series, so as a result of our analysis we obtain the vanishing of certain integrals

involving theta functions. The details of this will be worked out in Section 4, after we have developed the duality theory in Section 3.

3. Duality

Let f and ϕ be as in Section 1. We define first a map $\psi: (0, 1) \rightarrow (0, 1)$ such that

$$f(x)f(\psi(x)) = \frac{1}{\lambda}. \quad (1)$$

(Recall that $f(1) = 1$, $f(0) = (1/\lambda)$.) As f is strictly increasing and ranges from $1/\lambda$ to 1, it is clear that ψ exists and is unique. Explicitly:

$$\psi(x) = f^{(-1)}\left(\frac{1}{\lambda f(x)}\right).$$

From the symmetric form of (1) it is apparent that $\psi^{(2)} = \text{id}$. Put further $I = \phi \circ \psi$ and $\phi^* = I^{(-1)} \circ \psi = \psi \circ \phi^{(-1)} \circ \psi$. We will refer to ϕ^* as the *adjoint* map (with respect to f). As, solving out, $\phi = I \circ \psi = \psi \circ \phi^{*(-1)} \circ \psi$ it is clear that ϕ plays a similar rôle to ϕ^* as ϕ^* plays to ϕ , with $I^{(-1)}$ taking the rôle of I , formally: $\phi^{**} = \phi$.

If $\phi^* = \phi$ we say that ϕ is *self-adjoint* (with respect to f). This is the same as to say that I is an involution, i.e. $I^{(2)} = \text{id}$.

In this case (ϕ self-adjoint) the coefficients in the expansions for f and ϕ are related by the formula

$$\lambda = -\frac{c_0 d_1}{c_1 d_0}. \quad (2)$$

This is seen by expanding. Notice that d_0, d_1 and c_0 are > 0 while $c_1 < 0$.

We rewrite now the functional equation (3) in Section 1, using (1), as follows:

$$M(x) = f(x)M(\phi(x)),$$

$$M(\psi(x)) = f(\psi(x))M(I(x)),$$

$$\lambda f(x)M((I \circ I^{(-1)} \circ \psi)(x)) = M(I(x)),$$

$$\lambda f(x)(M \circ I)(\phi^*(x)) = (M \circ I)(x).$$

In other words, $M \circ I$ plays the same rôle for the map ϕ^* as Z plays for ϕ . By a similar argument, $Z \circ I$ plays the same rôle for ϕ^* as M plays for ϕ . Formally

$$Z^* = M \circ I, \quad M^* = Z \circ I.$$

In particular, in the self-adjoint case ($\phi = \phi^*$) we conclude that

$$Z = M \circ I, M = Z \circ I.$$

Example 1. Consider the case of the generalized arithmetic-geometric mean, i.e.

$$\phi(x) = \frac{2\sqrt{x + (\lambda - 2)x}}{1 + (\lambda - 1)x}, f(x) = \frac{1 + (\lambda - 1)x}{\lambda}$$

as in Example 1 in Section 1. Then ψ is determined by the formula

$$(1 + (\lambda - 1)x)(1 + (\lambda - 1)\psi(x)) = \lambda,$$

whence

$$(\lambda - 1)\psi(x) = \frac{\lambda}{1 + (\lambda - 1)x} - 1 = \frac{(\lambda - 1)(1 - x)}{1 + (\lambda - 1)x}$$

or

$$\psi(x) = \frac{1 - x}{1 + (\lambda - 1)x}.$$

Thus we find

$$\begin{aligned} I(x) = \phi(\psi(x)) &= \frac{2 \cdot \sqrt{\frac{1 - x}{1 + (\lambda - 1)x}} + (\lambda - 2) \frac{1 - x}{1 + (\lambda - 1)x}}{1 + (\lambda - 1) \frac{1 - x}{1 + (\lambda - 1)x}} \\ &= \frac{2\sqrt{(1 - x)(1 + (\lambda - 1)x)} + (\lambda - 2)(1 - x)}{\lambda}. \end{aligned}$$

If $\lambda = 2$ this formula takes the shape $I(x) = \sqrt{1 - x^2}$, well-known in the theory of the arithmetic-geometric mean [8], and if $\lambda = 4$ it can be found in [4], in which case it reads

$$I(x) = \frac{\sqrt{(1 - x)(1 + 3x)} + 1 - x}{2}.$$

We claim now that (for any λ) ϕ is self-adjoint.

Proof. We have to show that $I^{(2)} = \text{id}$. Writing $y = I(x)$ it follows that

$$(\lambda y - (\lambda - 2)(1 - x))^2 = 4(1 - x)(1 + (\lambda - 1)x)$$

or, after some simplifications,

$$\lambda(x + y - 1)^2 - 4(x - 1)(y - 1) = 0.$$

Our claim follows from the symmetric form of this expression. \square

In this case

$$d_0 = d_1 = \frac{\lambda - 1}{\lambda}, c_0 = \frac{1}{4}, c_1 = -\frac{1}{4\lambda},$$

in full agreement with formula (2).

Return to the general situation and consider Böttcher’s equation for $\phi^{(-1)}$, viz. the identity (8) in Section 1:

$$(N(\phi(x)))^2 = N(x).$$

Now we may rewrite it as follows

$$N(I(x))^2 = N(\phi(\psi(x)))^2 = N(\psi(x)) = N((I \circ I^{(-1)} \circ \psi)(x)) = N(I(\phi^*(x)))$$

or

$$(N \circ I)(x)^2 = (N \circ I)(\phi^*(x)).$$

That is, $N \circ I$ satisfies Böttcher’s equation for the map ϕ^* . (Note however that, as I is interchanging 0 and 1, the initial condition is a different one.) In particular, if ϕ is self-adjoint $N \circ I$ satisfies Böttcher’s equation for ϕ itself. Denoting the latter’s (“uniformizing”) solution by N^* , we then have $N^* = (N \circ I)^2$. This can again be proved by expanding.

Remark 1. In the self-adjoint case the algorithm can be given an elegant three dimensional formulation (in the case of the algorithms of Gauss and Borchardt this is known, see [8] and [4] respectively). Introduce, besides the two variables a, b , a third variable c by putting $c = aI(b/a)$ or, in homogenous notation, $c = I(a, b)$. The variables a, b, c are thus *not* dependent. Then we may consider the map

$$T: a, b, c \mapsto a' = f(a, b), b' = g(a, b), c' = h(a, b)$$

where $h(a, b) = f(a, b)\psi(a/b)$ (while $g(a, b) = f(a, b)\phi(a/b)$). Put further $R(a, b, c) = (a, c, b)$ (reflexion about the last two arguments). Then, on the one hand,

$$T: \begin{pmatrix} 1 \\ x \\ I(x) \end{pmatrix} \mapsto f(x) \begin{pmatrix} 1 \\ \phi(x) \\ \psi(x) \end{pmatrix},$$

and, on the other hand, as $I \circ \phi = \psi$, using this computation and formula (1)

$$T: f(x) \begin{pmatrix} 1 \\ \psi(x) \\ \phi(x) \end{pmatrix} \mapsto f(x) f(\psi(x)) \begin{pmatrix} 1 \\ \phi(\psi(x)) \\ \psi(\psi(x)) \end{pmatrix} = \frac{1}{\lambda} \begin{pmatrix} 1 \\ I(x) \\ x \end{pmatrix}.$$

Therefore we have

$$T \circ R \circ T = \frac{1}{\lambda} \cdot R$$

or, equivalently,

$$T^{(-1)} = \lambda R \circ T \circ R.$$

4. The algorithms of Gauss and Borchardt revisited and some other examples

We return, as promised at the end of Section 2, to the algorithms of Gauss and Borchardt and wish to express, in these cases, the conditions in (2) of that Section in terms of theta functions.

We begin with the arithmetic-geometric mean (Example 1 in the Introduction with $\lambda = 2$). In this case the relevant fact about theta functions is the set of formulae (see [3, p. 34]; we use the notation for theta functions of that book)

$$\frac{\theta_3^2(q) + \theta_4^2(q)}{2} = \theta_3(q^2), \quad \theta_3(q)\theta_4(q) = \theta_4(q^2). \tag{1}$$

Dividing these two equations with each other and writing

$$\chi(q) \stackrel{\text{def}}{=} \frac{\theta_4^2(q)}{\theta_3^2(q)},$$

we obtain

$$\frac{2\sqrt{\chi(q)}}{1 + \chi(q)} = \chi(q^2),$$

$$\phi(\chi(q)) = \chi(q^2),$$

which is the “inverse” form of Böttcher’s equation for the map ϕ . That is, it is formally equivalent to

$$\chi^{(-1)}(x)^2 = \chi^{(-1)}(\phi(x)).$$

In particular, we conclude that

$$\chi = (N^*)^{(-1)},$$

where N^* stands for the latter's ("uniformizing") solution.

Look now at the formula

$$\int_0^1 \left(\log \frac{1}{N(x)} \right)^{(2\pi i n / \log 2)} d \log \frac{1+x}{2} = 0 \quad (n \in \mathbf{N}),$$

which is just (2) in Section 2

$$\left(f(x) = \frac{1+x}{2} \text{ in this case!} \right).$$

In this formula we make the change of variable $q = N(x)^2$. As $N = (N^* \circ I)^{1/2}$ this is the same as $q = N^*(Ix)$ so that (cf. [3, p. 35])

$$x = I(\chi(q)) = \sqrt{1 - \frac{\theta_4^4(q)}{\theta_3^4(q)}} = \frac{\theta_2^2(q)}{\theta_3^2(q)}.$$

It follows that

$$\int_0^1 (\log 1/q)^{(2\pi i n / \log 2)} d \log \left(\frac{1 + (\theta_2^2(q)/\theta_3^2(q))}{2} \right) = 0 \quad (n \in \mathbf{N}).$$

Remark 1. We note also that the "asymptotic" formula

$$M = \frac{\pi}{2} \frac{M \circ I}{\log(1/N)}$$

can be written

$$N = e^{(\pi/2)(M/M \circ I)}.$$

Using $N = (N^* \circ I)^{1/2}$ once more, we obtain Gauss's famous formula

$$q = e^{-\pi(M(I(x))/M(x))}.$$

Consider now Borchardt's generalization ($l=4$). In this case we have, as a generalization of (1), see [4],

$$\frac{\mathcal{L}^2(q) + 3\mathcal{M}^2(q)}{4} = \mathcal{L}^2(q^2), \quad \frac{\mathcal{L}(q)\mathcal{M}(q) + \mathcal{M}^2(q)}{2} = \mathcal{M}^2(q^2),$$

where we have written, still following [4],

$$\begin{aligned} \mathcal{L}(q) &\stackrel{\text{def}}{=} \theta_3(q)\theta_3(q^3) + \theta_2(q)\theta_2(q^3), \\ \mathcal{M}(q) &\stackrel{\text{def}}{=} \theta_3(q)\theta_3(q^3) + \theta_2(q)\theta_2(q^3) (= \theta_4(q)\theta_4(q^3)). \end{aligned}$$

Putting

$$\chi(q) = \frac{\mathcal{M}(q)}{\mathcal{L}(q)}$$

we have again

$$\chi = (N^*)^{(-1)},$$

so we can make a similar transformation of formula (2), Section 2 in this case. However, owing to the complicated form of I now we have not been able to write down such an explicit result as in Gauss's case.

Another example of an algorithm to which our "theory" applies is the cubic scheme encountered by the Borweins in another remarkable paper [5]:

$$a, b \mapsto a' = \frac{a + 2b}{3}, \quad b' = \sqrt[3]{b \left(\frac{a^2 + ab + b^2}{3} \right)};$$

it is cubic in the sense that the fixed points 0 and 1 are cubic, not quadratic repelling respectively attracting, which of course matters only little. In this case

$$\phi(x) = \frac{\sqrt[3]{9x(1+x+x^2)}}{1+2x},$$

$$\psi(x) = \frac{1-x}{1+2x},$$

$$I(x) = \sqrt[3]{1-x^3}.$$

From the last expression we readily see that it is question of a selfadjoint situation in our sense; it is related to the Fermat curve $x^3 + y^3 = 1$ rather than the circle $x^2 + y^2 = 1$ in the classical case of the arithmetic-geometric mean.

In the same paper [5] there appears another interesting example, viz. the iterative scheme

$$a, b \mapsto a' = \frac{a+3b}{4}, b' = \sqrt{b \frac{a+b}{2}},$$

which is fully analyzed there—the limit is given in terms of the hypergeometric function—and at least formally fits into the present scheme.

Finally, in the paper [4], which we have been quoting repeatedly, there occurs however also the iteration

$$a, b \mapsto a' = \frac{a+2b-\sqrt{ab}}{2}, b' = \frac{\sqrt{ab+b}}{2},$$

to which our “theory” does *not* apply in full. This is because in this case

$$f(x) = \frac{1+2x-\sqrt{x}}{2}$$

is, unfortunately, not an increasing function in the whole interval (0, 1). In particular, it is meaningless to speak about self-adjointness.

5. The generalized arithmetic-geometric mean for complex values

We return to the generalized arithmetic-geometric mean, viz. the iterative scheme (cf. formula (5) in the Introduction)

$$a, b \mapsto a' = \frac{a+(\lambda-1)b}{\lambda}, b' = \frac{\sqrt{ab+(\lambda-2)b}}{\lambda}, \quad (1)$$

but now we allow a and b , hitherto assumed to be real positive, to take complex values. In addition, we also assume that $\lambda \geq 2$. Because of the two-valuedness of the square root, the interpretation of (1) is now ambiguous.

Definition. Let b' and b'^* be the two different determinations of b' . We say that b' is the *good choice* and b'^* the *second choice* if

$$|a' - b'| \leq |a' - b'^*|.$$

We have

$$b' = \frac{2\sqrt{ab+(\lambda-2)b}}{\lambda}, b'^* = \frac{-2\sqrt{ab+(\lambda-2)b}}{\lambda}$$

with a fixed interpretation of $\sqrt{}$. Therefore

$$a' - b' = \frac{a + b - 2\sqrt{ab}}{\lambda}, \quad a' - b'^* = \frac{a + b + 2\sqrt{ab}}{\lambda},$$

whence

$$(a' - b')(a' - b'^*) = \frac{(a + b)^2 - 4ab}{\lambda^2} = \frac{(a - b)^2}{\lambda^2}.$$

It follows that if b' is a good choice then

$$|a' - b'| \leq \frac{1}{\lambda} |a - b|. \quad (2)$$

Note also that

$$a' - a = \frac{\lambda - 1}{\lambda} (b - a). \quad (3)$$

Now iterate the map in (1) with given initial data, taking alternatively good and second choices according to some prescribed law. It follows immediately from (2) and (3) that we have:

Theorem 1. *If we make at most a finite number of second choices then the algorithm is convergent.*

Consider now

$$\begin{aligned} a'' &= \frac{a' + (\lambda - 1)b'}{\lambda} = \frac{a' - b'^* + (\lambda - 1)b' + b'^*}{\lambda} \\ &= \frac{a' - b'^*}{\lambda} + \frac{(\lambda - 1)b' + b'^*}{\lambda}. \end{aligned}$$

If b' is a *second* choice (*sic!*) then in view of (2) we can estimate the first term here by

$$\frac{1}{\lambda^2} |a - b|. \quad (4)$$

On the other hand, we can write

$$\begin{aligned}
 (\lambda-1)b' + b'^* &= \frac{1}{\lambda} [(\lambda-1)2\sqrt{ab} + (\lambda-1)(\lambda-2)b - 2\sqrt{ab} + (\lambda-2)b] \\
 &= \frac{\lambda-2}{\lambda} [2\sqrt{ab} + \lambda b].
 \end{aligned}$$

Set $\omega = \max(|a|, |b|)$ and let ω' , ω'' and ω''' have an analogous meaning. As $\lambda \geq 2$ by hypothesis, it is clear that $\omega \geq \omega' \geq \omega''$. It follows that

$$|(\lambda-1)b' + b'^*| \leq \frac{\lambda-2}{\lambda} (\lambda+2)\omega = \frac{\lambda^2-4}{\lambda} \omega$$

and, as $|a-b| \leq 2\omega$, together with (4) this gives

$$|a''| \leq \frac{1}{\lambda^2} (2 + \lambda^2 - 4)\omega = \frac{\lambda^2-2}{\lambda^2} \omega \equiv \Lambda\omega \quad (5)$$

with $\Lambda < 1$. It is likewise clear that

$$|b''| \leq \omega. \quad (6)$$

From (5) and (6) it follows that

$$\omega''' \leq \Lambda_1 \omega \quad (7)$$

with $\Lambda_1 < 1$. Indeed, we can take

$$\Lambda_1 = \frac{2\Lambda + \lambda - 2}{\lambda} = 1 - \frac{4}{\lambda^3}.$$

Using (7) we readily get

Theorem 2. *If we make infinitely many second choices then the algorithm converges also but the limit is always zero. \square*

Remark 1. For $\lambda=2$ these theorems are in essence due to Gauss [8]. Our treatment is copied on Cox [7]. Recently, an alternative proof in Gauss's case, based on monodromy, was given by Nishiwada [9]. In the case $\lambda > 2$ our result is not entirely exhaustive; it remains to determine when the limit is nonzero in the first case (Theorem 1). There remains also to connect the various values of it, what was of course Gauss's main objective in the classical case. But, as we have already pointed out in the Introduction, it is hardly likely that this can be done more explicitly unless $\lambda=4$.

However, one can always ask weaker questions. For instance, does the limit set have other points of accumulation than the origin?

REFERENCES

1. J. ARAZY, T. CLAESSON, S. JANSON and J. PEETRE, Means and their iterations, in *Proc. of the Nineteenth Nordic Congress of Mathematicians, Reykjavik, 1984* (Jon R. Stefanson (ed.), Icelandic Math. Soc., Science Inst., University of Iceland, Reykjavik, 1985), 191–212.
2. C.-W. BORCHARDT, Ueber das arithmetisch-geometrische Mittel aus vier Elementen, *Monatsh. Akad. Wiss. Berlin* (1876), 611–621.
3. J. M. BORWEIN and P. M. BORWEIN, *Pi and the AGM—A Study in Analytic Number Theory and Computational Complexity* (John Wiley, New York, 1987).
4. J. M. BORWEIN and P. M. BORWEIN, On the mean iteration $(a, b) \leftarrow ((a + 3b/4), (\sqrt{ab} + b/2))$, *Math. Comp.* **53** (1989), 311–326.
5. J. M. BORWEIN and P. M. BORWEIN, A cubic counterpart of Jacobi's identity and the AGM, *Trans. Amer. Math. Soc.* **323** (1991), 691–701.
6. J. M. BORWEIN and P. M. BORWEIN, *Quadratic Mean Iterations*, in preparation.
7. D. A. COX, The arithmetic-geometric mean of Gauss, *Ensign. Math.* **30** (1984), 275–330.
8. C. F. GAUSS, *Bestimmung der Anziehung eines elliptischen Ringes. Nachlass zur Theorie des arithmetisch-geometrischen Mittels*, (Übersetz von Dr Harald Geppert, Ostwalds Klassiker der Exakten Wissenschaften **225**, Akademische Verlagsgesellschaft, Leipzig, 1927).
9. K. NISHIWADA, A holomorphic structure of the arithmetic-geometric mean of Gauss, *Proc. Japan Acad. Ser. A* **64** (1988), 322–324.
10. J. PEETRE, Some unsolved problems, in *Colloquia Mathematica Societatis Janos Bolyai 49*, (Alfred Haar Memorial Conference, Budapest (Hungary), 1985, North-Holland, Amsterdam, 1986), 711–735.
11. J. PEETRE, Generalizing the arithmetic-geometric mean—a hapless computer experiment, *Internat. J. Math. Sci.* **12** (1989), 235–246.

MATEMATISKA INSTITUTIONEN
STOCKHOLMS UNIVERSITET
Box 6701
S-113 85 STOCKHOLM
SWEDEN