

Rational Approximants Defined from Double Power Series

By J. S. R. Chisholm

Abstract. Rational approximants are defined from double power series in variables x and y , and it is shown that these approximants have the following properties: (i) they possess symmetry between x and y ; (ii) they are in general unique; (iii) if $x = 0$ or $y = 0$, they reduce to diagonal Padé approximants; (iv) their definition is invariant under the group of transformations $x = Au/(1 - Bu)$, $y = Av/(1 - Cv)$ with $A \neq 0$; (v) an approximant formed from the reciprocal series is the reciprocal of the corresponding original approximant. Possible variations, extensions and generalisations of these results are discussed.

1. Introduction. When several terms of a power series

$$(1.1) \quad f(z) = \sum_{\gamma=0}^{\infty} c_{\gamma} z^{\gamma}$$

are known, forming Padé approximants from the series is frequently an effective way of approximating the function $f(z)$ represented by the series expansion. References to several general works on the subject are given ([1], [2], [3], [4]). The (m, n) approximant

$$(1.2) \quad f_{m,n}(z) \equiv \sum_{\alpha=0}^m a_{\alpha} z^{\alpha} / \sum_{\beta=0}^n b_{\beta} z^{\beta}$$

is defined by the formal identity

$$(1.3) \quad \left[\sum_{\beta=0}^n b_{\beta} z^{\beta} \right] \left[\sum_{\gamma=0}^{\infty} c_{\gamma} z^{\gamma} \right] = \sum_{\alpha=0}^m a_{\alpha} z^{\alpha} + o(z^{m+n}).$$

In general, the $(m + n + 1)$ linear equations obtained by equating powers of $1, z, \dots, z^{m+n}$ in (1.3) define uniquely the ratios of $\{a_{\alpha}\}$ and $\{b_{\beta}\}$, and hence $f_{m,n}(z)$. The diagonal approximants $\{f_{m,m}(z)\}$ have proved to be particularly powerful; their importance stems from certain invariance properties [5], the most important of which are the following:

(1) If we substitute

$$(1.4) \quad z = Aw/(1 - Bw) \quad (A \neq 0)$$

in (1.1) and expand the denominator factors $(1 - Bw)^{-\gamma}$ by the binomial theorem to give a new series

$$(1.5) \quad g(w) = \sum_{\epsilon=0}^{\infty} e_{\epsilon} w^{\epsilon},$$

Received October 10, 1972, revised November 27, 1972.

AMS (MOS) subject classifications (1970). Primary 30A82, 41A20; Secondary 30A14, 30A22, 41A63.

Key words and phrases. Approximation theory, rational approximation, double series, Padé approximants, invariance properties.

Copyright © 1973, American Mathematical Society

then the $[m/m]$ Padé approximant $g_{m,m}(w)$ formed from (1.5) obeys

$$(1.6) \quad g_{m,m}(w) \equiv f_{m,m}(Aw/(1 - Bw)).$$

In other words, the formation of diagonal approximants is invariant under the group of homographic transformations defined by (1.4).

(2) If the reciprocal of series (1.1) is

$$(1.7) \quad f^{-1}(z) = \sum_{\delta=0}^{\infty} d_{\delta} z^{\delta},$$

then the diagonal approximants to $f^{-1}(z)$ are given by

$$(1.8) \quad f_{m,m}^{-1}(z) = [f_{m,m}(z)]^{-1}.$$

If we try to define analogous approximants from a double power series

$$(1.9) \quad \sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} x^{\alpha} y^{\beta}$$

in variables x and y , it seems natural to assume that the numerators and denominators contain all terms up to a given order in x and y ; for example, a $[2/2]$ approximant of this type would contain terms $1, x, y, x^2, xy$ and y^2 in both numerator and denominator. It is easy to see that no formal equation analogous to (1.3) gives the correct number of linear equations to determine the ratios of the coefficients in the approximants. Gammel [6] has suggested using a minimisation procedure to determine the "best possible" values of the coefficients, but it is not clear that Gammel's approximants share the attractive properties of Padé approximants.

The 2-variable approximants defined in the next section are generalisations of diagonal Padé approximants, and will be shown to have the following properties:

- (i) The approximants are symmetrical between x and y .
- (ii) The number of linear equations is equal to the number of ratios of coefficients to be determined, so that, in general, a unique approximant is defined.
- (iii) If $x = 0$ or $y = 0$, the approximants become diagonal Padé approximants in the other variable.
- (iv) The formation of the approximants from (1.9) is invariant under all transformations of the group

$$(1.10) \quad x = \frac{Au}{1 - Bu}, \quad y = \frac{Av}{1 - Cv} \quad (A \neq 0).$$

(v) If

$$(1.11) \quad \sum_{\gamma, \delta=0}^{\infty} d_{\gamma\delta} x^{\gamma} y^{\delta}$$

is the reciprocal of the series (1.9), then the reciprocal of an approximant defined from (1.9) is equal to the corresponding approximant defined from (1.11).

Further, we shall see that these are the only approximants satisfying all these properties, and that the choice of the formal equality is a natural one, given the conditions (i)–(v). In Section 4, we shall discuss possible variants and generalisations of the new approximants.

2. Definition of the Diagonal Approximants. The approximants are defined to have a given maximum power in each variable, rather than to have given total maximum power. The $[m/m]$ approximant is thus of the form

$$(2.1) \quad f_{m,m}(x, y) \equiv \sum_{\mu, \nu=0}^m a_{\mu\nu} x^\mu y^\nu / \sum_{\sigma, \tau=0}^m b_{\sigma\tau} x^\sigma y^\tau.$$

Without loss of generality, we can normalise the series (1.9) by taking

$$(2.2) \quad c_{00} = 1;$$

in general, this will allow us to choose

$$(2.3) \quad a_{00} = b_{00} = 1.$$

In this paper, we shall not consider series (1.9) which are not “normal” in the Padé sense ([1], [4]). The $[1/1]$ approximant, for example, is of the form

$$(2.4) \quad f_{1,1}(x, y) = \frac{1 + a_{10}x + a_{01}y + a_{11}xy}{1 + b_{10}x + b_{01}y + b_{11}xy},$$

with numerator and denominator both bilinear in x and y . The number of coefficients to be determined in $f_{m,m}$ is, remembering (2.3),

$$(2.5) \quad 2[(m + 1)^2 - 1] = 2m^2 + 4m.$$

The most natural extension of (1.3) is the formal equality

$$(2.6) \quad \left[\sum_{\sigma, \tau=0}^m b_{\sigma\tau} x^\sigma y^\tau \right] \left[\sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} x^\alpha y^\beta \right] = \sum_{\mu, \nu=0}^m a_{\mu\nu} x^\mu y^\nu + o(x^\gamma y^{2m-\gamma}).$$

The last term in (2.6) indicates that coefficients of all terms of total order less than $(2m + 1)$ are equated. Since (2.2) and (2.3) ensure that the zero order terms are equal, the number of equations is

$$(2.7) \quad \sum_{r=2}^{2m+1} r = 2m^2 + 3m,$$

which is m less than the number of unknowns (2.5). We therefore require m further equations; it is reasonable to obtain these conditions by equating coefficients in (2.6) of terms of the type

$$(2.8) \quad x^\gamma y^{2m+1-\gamma},$$

of total order $(2m + 1)$; there are $(2m + 2)$ coefficients of this form. We first note that it is inconsistent to equate coefficients of x^{2m+1} and y^{2m+1} in (2.6); this is most easily seen by considering (2.6) with $y = 0$. It is clear that the equation then reduces to the usual definition of an $[m/m]$ Padé approximant in the variable x , formed by equating coefficients up to order x^{2m} . In general, one cannot also match the coefficients of x^{2m+1} . We can therefore only consider the $2m$ terms of type (2.8) with $\gamma = 1, 2, \dots, 2m$. It seems that the only way of obtaining m linear conditions, symmetrical in x and y , is to equate the *sums of coefficients* of the m pairs of terms

$$(2.9) \quad x^\gamma y^{2m+1-\gamma}, \quad x^{2m+1-\gamma} y^\gamma \quad (\gamma = 1, 2, \dots, m).$$

The justification of this choice is that it defines approximants satisfying all the criteria (i)–(v). In Section 4 we shall discuss other possible choices of linear equations.

If we define

$$(2.10) \quad a_{\gamma\delta} = b_{\gamma\delta} \equiv 0 \quad (\gamma > m \text{ or } \delta > m),$$

then the linear equations arising from (2.6) and from the pairs (2.9) are

$$(2.11) \quad \sum_{\sigma=0}^{\gamma} \sum_{\tau=0}^{\delta} b_{\sigma\tau} c_{\gamma-\sigma, \delta-\tau} = a_{\gamma\delta} \quad (\gamma, \delta = 0, 1, 2, \dots, 2m; 1 \leq \gamma + \delta \leq 2m)$$

and

$$(2.12) \quad \sum_{\sigma=0}^{\gamma} \sum_{\tau=0}^{\delta} (b_{\sigma\tau} c_{\gamma-\sigma, \delta-\tau} + b_{\tau\sigma} c_{\delta-\tau, \gamma-\sigma}) = 0$$

$$(\gamma = 1, 2, \dots, m; \gamma + \delta = 2m + 1).$$

Of the Eqs. (2.11), $m(m+2)$ define the coefficients $\{a_{\gamma\delta}\}$ ($\gamma, \delta = 0, 1, \dots, m$), excluding a_{00} . The remaining $m(m+1)$ equations, together with the m equations (2.12), define the $m(m+2)$ coefficients $\{b_{\sigma\tau}\}$, remembering that $b_{00} = 1$.

Equation (2.6) is equivalent to the set of Eqs. (2.11). To indicate the additional equalities (2.12), we amend (2.6) to give

$$(2.13) \quad \left[\sum_{\sigma, \tau=0}^m b_{\sigma\tau} x^{\sigma} y^{\tau} \right] \left[\sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} x^{\alpha} y^{\beta} \right] = \sum_{\mu, \nu=0}^m a_{\mu\nu} x^{\mu} y^{\nu} + o[x^{\gamma} y^{2m-\gamma}, S(x^{\gamma} y^{2m+1-\gamma})],$$

the symbol “ S ” denoting symmetrisation.

To exemplify the definition, we shall write down the equations determining $f_{1,1}$. Equating coefficients of x, y, xy, x^2 and y^2 in (2.13) gives, remembering (2.2) and (2.3),

$$(2.14) \quad \begin{aligned} b_{10} + c_{10} &= a_{10}, \\ b_{01} + c_{01} &= a_{01}, \\ b_{11} + b_{10}c_{01} + b_{01}c_{10} + c_{11} &= a_{11}, \\ c_{10}b_{10} + c_{20} &= 0, \\ c_{01}b_{01} + c_{02} &= 0. \end{aligned}$$

These are the five equations (2.11). There is only one equation in the set (2.12), given by equating the sum of coefficients of x^2y, xy^2 :

$$(2.15) \quad (c_{21} + c_{12}) + (c_{11} + c_{02})b_{10} + (c_{11} + c_{20})b_{01} + (c_{01} + c_{10})b_{11} = 0.$$

We note that the equations defining a_{10}, a_{01}, b_{10} and b_{01} are just the equations defining the two $[1/1]$ Padé approximants given by putting $y = 0$ or $x = 0$ in (2.6). We have already noted that this property (iii) follows in general from (2.6) or (2.13) by putting $x = 0$ or $y = 0$.

We further note that in (2.14) and (2.15), the highest order coefficients c_{12} and c_{21} occur only in the combination $(c_{12} + c_{21})$. Generally, pairs of highest order coefficients $(c_{\alpha, 2m+1-\alpha}, c_{2m+1-\alpha, \alpha})$ occur only in the combinations

$$(2.16) \quad (c_{\alpha, 2m+1-\alpha} + c_{2m+1-\alpha, \alpha})$$

in Eqs. (2.12), and do not occur in Eqs. (2.11). This fact will be important in establishing property (v) of the approximants.

3. Invariance Properties. Of the properties (i)–(v) enumerated at the end of Section 1, (i) and (ii) are satisfied directly by definition of the approximants, and (iii) was established in Section 2. We shall now establish the important invariance properties (iv) and (v).

The first step in establishing (iv) is to note that the substitution of (1.10) into the function (2.1) gives the rational function

$$\begin{aligned}
 (3.1) \quad g_{m,m}(u, v) &\equiv \frac{\sum_{\mu, \nu=0}^m a_{\mu\nu} (Au)^\mu (1 - Bu)^{m-\mu} (Av)^\nu (1 - Cv)^{m-\nu}}{\sum_{\sigma, \tau=0}^m b_{\sigma\tau} (Au)^\sigma (1 - Bu)^{m-\sigma} (Av)^\tau (1 - Cv)^{m-\tau}} \\
 &\equiv \frac{\sum_{\mu, \nu=0}^m a'_{\mu\nu} u^\mu v^\nu}{\sum_{\sigma, \tau=0}^m b'_{\sigma\tau} u^\sigma v^\tau},
 \end{aligned}$$

say. The highest powers of u and v in both the numerator and the denominator in (3.1) are u^m and v^m ; $g_{m,m}$ is thus a function of the same form as $f_{m,m}$. If it can be shown that $g_{m,m}$ obeys the equation analogous to (2.11) and (2.12), but in variables u and v , it follows that $g_{m,m}$ is the $[m/m]$ approximant in these variables. The series from which this $[m/m]$ approximant is defined is obtained by substituting (1.10) into (1.9), giving

$$(3.2) \quad \sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} \left(\frac{Au}{1 - Bu} \right)^\alpha \left(\frac{Av}{1 - Cv} \right)^\beta,$$

and then expanding the factors $(1 - Bu)^{-\alpha}$ and $(1 - Cv)^{-\beta}$ ($\alpha, \beta = 1, 2, \dots$) by the binomial theorem; this procedure defines as many terms as desired of a formal power series in u, v . We denote this power series by

$$(3.3) \quad \sum_{\xi, \eta=0}^{\infty} c'_{\xi\eta} u^\xi v^\eta.$$

It is clear that a particular coefficient $c_{\alpha\beta}$ in (3.2) only contributes to coefficients $c'_{\xi\eta}$ in (3.3) which have $\xi \geq \alpha$ and $\eta \geq \beta$.

If we define a double series

$$(3.4) \quad \sum_{\mu, \nu=0}^{\infty} r_{\mu\nu} x^\mu y^\nu \equiv \sum_{\mu, \nu=0}^m a_{\mu\nu} x^\mu y^\nu - \left[\sum_{\sigma, \tau=0}^m b_{\sigma\tau} x^\sigma y^\tau \right] \left[\sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} x^\alpha y^\beta \right],$$

Eqs. (2.11) and (2.12) can be expressed as

$$(3.5) \quad r_{\mu\nu} = 0 \quad (\mu + \nu \leq 2m)$$

and

$$(3.6) \quad r_{\mu, 2m+1-\mu} + r_{2m+1-\mu, \mu} = 0, \quad \mu = (1, \dots, m).$$

Remembering (3.1), we see that multiplying (3.4) by $(1 - Bu)^m(1 - Cv)^m$ gives

$$\begin{aligned}
& \sum_{\mu, \nu=0}^m a_{\mu\nu} (Au)^\mu (1 - Bu)^{m-\mu} (Av)^\nu (1 - Cv)^{m-\nu} \\
& + \left[\sum_{\sigma, \tau=0}^m b_{\sigma\tau} (Au)^\sigma (1 - Bu)^{m-\sigma} (Av)^\tau (1 - Cv)^{m-\tau} \right] \\
(3.7) \quad & \cdot \left[\sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} \left(\frac{Au}{1 - Bu} \right)^\alpha \left(\frac{Av}{1 - Cv} \right)^\beta \right] \\
& = \sum_{\mu, \nu=0}^m a'_{\mu\nu} u^\mu v^\nu - \left[\sum_{\sigma, \tau=0}^m b'_{\sigma\tau} u^\sigma v^\tau \right] \left[\sum_{\alpha, \beta=0}^{\infty} c'_{\alpha\beta} u^\alpha v^\beta \right] \\
& \equiv \sum_{\mu, \nu=0}^{\infty} r'_{\mu\nu} u^\mu v^\nu,
\end{aligned}$$

say. By considering the coefficient of $u^\mu v^\nu$ in Eq. (3.7), it is not difficult to show that

$$(3.8) \quad r'_{\mu\nu} = \sum_{\gamma=0}^{\mu} \sum_{\delta=0}^{\nu} \binom{m - \mu + \gamma}{\gamma} \binom{m - \nu + \delta}{\delta} A^{\mu+\nu-\gamma-\delta} (-B)^\gamma (-C)^\delta r_{\mu-\gamma, \nu-\delta}$$

($\mu \leq m, \nu \leq m$)

and that $r'_{\mu\nu} = 0$ ($\mu > m$ or $\nu > m, \mu + \nu \leq 2m$). Using (3.5), it follows that

$$(3.9) \quad r'_{\mu\nu} = 0 \quad (\mu + \nu \leq 2m).$$

Similarly, it follows that

$$(3.10) \quad r'_{\mu\nu} = A^{2m+1} r_{\mu\nu} \quad (\mu + \nu = 2m + 1).$$

Thus, (3.6) implies that

$$(3.11) \quad r'_{\mu, 2m+1-\mu} + r'_{2m+1-\mu, \mu} = 0 \quad (\mu = 1, \dots, m).$$

Comparison of (3.9) and (3.11) with (3.5) and (3.6) establishes the invariance of these equations under transformations of the form (1.10).

We now establish property (v) of Section 1. The reciprocal (1.11) of the double series (1.9) satisfies the formal equality

$$(3.12) \quad \left[\sum_{\alpha, \beta=0}^{\infty} c_{\alpha\beta} x^\alpha y^\beta \right] \left[\sum_{\gamma, \delta=0}^{\infty} d_{\gamma\delta} x^\gamma y^\delta \right] \equiv 1,$$

so that the coefficients $\{d_{\gamma\delta}\}$ satisfy the "triangular" set of equations

$$(3.13) \quad \sum_{\gamma=0}^{\mu} \sum_{\delta=0}^{\nu} c_{\mu-\gamma, \nu-\delta} d_{\gamma\delta} = 0 \quad (\mu + \nu \geq 1).$$

If we formally multiply (3.4) by the series (1.11), we obtain

$$\begin{aligned}
& - \left[\sum_{\gamma, \delta=0}^{\infty} d_{\gamma\delta} x^\gamma y^\delta \right] \left[\sum_{\mu, \nu=0}^{\infty} r_{\mu\nu} x^\mu y^\nu \right] \\
(3.14) \quad & = \sum_{\sigma, \tau=0}^m b_{\sigma\tau} x^\sigma y^\tau - \left[\sum_{\mu, \nu=0}^m a_{\mu\nu} x^\mu y^\nu \right] \left[\sum_{\gamma, \delta=0}^{\infty} d_{\gamma\delta} x^\gamma y^\delta \right] \\
& \equiv \sum_{\sigma, \tau=0}^{\infty} r''_{\sigma\tau} x^\sigma y^\tau.
\end{aligned}$$

Equations (3.5) therefore imply that

$$(3.15) \quad r''_{\mu\nu} = 0 \quad (\mu + \nu \leq 2m)$$

and that

$$r''_{\mu\nu} = -r_{\mu\nu} \quad (\mu + \nu = 2m + 1).$$

So (3.6) gives

$$(3.16) \quad r''_{\mu, 2m+1-\mu} + r''_{2m+1-\mu, \mu} = 0.$$

Equations (3.14), (3.15) and (3.16) show that the $[m/m]$ approximant to the reciprocal series (1.11) is $[f_{m,m}]^{-1}$.

We have therefore shown that the two-variable approximants satisfy all the properties (i)–(v). Property (iii) ensures that the approximants will be as powerful as Padé approximants in regions near $x = 0$ and $y = 0$. We expect, however, that property (iv) will ensure the usefulness of the approximants over large ranges of values of the complex variables x and y . The transformations of the group (1.10) do not allow relative changes of scale of x and y ; otherwise it allows all independent transformations of the corresponding Padé invariance groups. This implies that the use of variables (x, Dy) , with $D \neq 1$, will define different approximants, since the relative weighting of terms in (2.12) will change. In practice, therefore, we should try to choose D so that the coefficients in the double series (1.9) have a variation which is roughly symmetrical between x and y .

4. Variants and Generalisations. By relaxing one or more of the conditions (i)–(v), one can invent a number of variants of the approximants defined in Section 2. If the information we have about the coefficients $\{c_{r,s}\}$ is unsymmetrical between x and y , or if we require approximants giving a better representation in one variable than the other, we could use approximants which lack symmetry between x and y . Two ways of “breaking symmetry” are:

(a) Using different maximum powers of x and y in the numerator and denominator polynomials; they could, for example, be linear in x and quadratic in y .

(b) Making an unsymmetrical choice of coefficients which are equated. To give a simple example, we could replace (2.15) by

$$c_{12} + c_{02}b_{10} + c_{11}b_{01} + c_{10}b_{11},$$

using only the coefficients of xy^2 ; we do not then need to know c_{21} .

Another variant which could be considered is:

(c) Using different order numerator and denominator polynomials, analogous to nondiagonal Padé approximants.

The fact that the approximants are generalisations of diagonal Padé approximants suggests a wide variety of ways in which properties of Padé approximants might be extended. In particular, the following investigations suggest themselves:

(d) A study of the numerical accuracy of the approximants to particular functions of two variables.

(e) A study of the relation between the singularity structures of particular functions and their approximants.

(f) A search for algorithms.

- (g) Applications to double sequences and to sequences of series.
- (h) Studying the relation to the moment problem in two variables, and looking for an analogue to series of Stieltjes.
- (i) Extension of the theory to include double series whose coefficients $\{c_{\alpha\beta}\}$ are square matrices.
- (j) Investigating the possibility of defining generalised approximants, analogous, for example, to Gammel-Baker approximants [7].
- (k) Defining analogous approximants to series in three or more variables.

The program of work involved in these suggested investigations is very large; one must therefore be selective in choosing a line of investigation. For this reason, we have concentrated in this paper upon rational approximants which possess all the properties (i)–(v). The most immediate task seems to be a numerical investigation of their accuracy and their singularity structure.

Acknowledgement. I would like to thank Dr. P. R. Graves-Morris and Dr. R. Hughes Jones for carefully reading this manuscript, and also Dr. Hughes Jones for correcting an error in defining the group of transformations (1.10). I am grateful for suggestions of the referee, which considerably improved the presentation of the proofs in Section 3. My thanks are also due to Mrs. Bateman for typing and re-typing the manuscript with speed, care and accuracy.

After the revised edition of this paper had been submitted, the author received a copy of work by C. Lutterödt which also described two-variable approximants. Lutterödt's choice of approximants is different from those studied in this paper; in particular, they do not reduce to Padé approximants when one variable is equated to zero.

Mathematical Institute
University of Kent
Canterbury, Kent, England

1. *Proceedings of the Canterbury Summer School on Padé Approximants and Their Applications*, P. R. Graves-Morris (Editor), Institute of Physics, 1973.
2. *The Padé Approximant in Theoretical Physics*, G. A. Baker, Jr. and J. L. Gammel (Editors), Academic Press, New York, 1970, pp. 20–29.
3. O. PERRON, *Die Lehre von den Kettenbrüchen*, Dritte, Verbesserte und erweiterte Aufl. Band II, Analytisch-funktionen-theoretische Kettenbrüche, Teubner Verlagsgesellschaft, Stuttgart, 1957. MR **19**, 25.
4. H. S. WALL, *Analytic Theory of Continued Fractions*, Van Nostrand, Princeton, N.J., 1948. MR **10**, 32.
5. G. A. BAKER, JR., J. L. GAMMEL & J. G. WILLS, "An investigation of the applicability of the Padé approximant method," *J. Math. Anal. Appl.*, v. 2, 1961, pp. 405–418. MR **23** #B3125.
6. J. L. GAMMEL, *Proceedings of the 1972 Canterbury International Conference on Padé Approximants and Their Applications*, P. R. Graves-Morris (Editor), Academic Press, New York, 1973.