ORTHOGONAL POLYNOMIALS AND HYPERGEOMETRIC SERIES

A. VAN DER SLUIS

Introduction. In Part I of this paper we present a theory of Padé-approximants for Laurent series, and discuss their relation to orthogonal polynomials. For earlier results in this direction we may refer to **(1; 7; 8)**. It is also indicated how this theory can be extended, for example, to matrix polynomials.

In order to derive certain special types of orthogonal polynomials we need explicit expressions for Padé-approximants. In Part II we generalize a result of Padé **(5)**, giving such expressions in the hypergeometric case. The resulting polynomials are the classical ones and basic analogues of them. Concerning these analogues see also Hahn **(2)**.

In the final part it is proved that under a much more natural and apparently less restrictive condition no more general polynomials result than those obtained in Part II.

PART I

1. Orthogonal polynomials. Our definition of orthogonality will be similar to the generalized definition of Krall **(3)**. Suppose we are given a sequence $r_0, r_1, r_2, \ldots$, in a field R, such that each set of equations

$$(1.1)\qquad \begin{cases} q_0 r_0 + q_1 r_1 + \ldots + q_m r_m = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ q_0 r_{m-1} + q_1 r_m + \ldots + q_m r_{2m-1} = 0 \end{cases} \qquad m = 1, 2, 3, \ldots,$$

has exactly one solution with $q_m = 1$ in R (in that case the sequence $\{r_n\}$ will be called *regular*). We then define a *moment operator* Ω operating on polynomials as follows

$$(1.2)\qquad \Omega \sum_{\mu=0}^{m} p_\mu x^\mu = \sum_{\mu=0}^{m} p_\mu r_\mu,$$

and the set of polynomials $Q_0(x)$, $Q_1(x)$, $Q_2(x), \ldots$, over R of respective degrees $0, 1, 2, \ldots$, is called *orthogonal* with respect to the sequence r_0, r_1, $r_2, \ldots$, if

$$(1.3)\qquad \Omega Q_m(x) x^n = 0 \qquad n < m,$$

Received June 27, 1957; in revised form February 14, 1958.

Parts I and II of this paper cover parts of a thesis submitted by the author to the University of Amsterdam (*General orthogonal polynomials*, Groningen, 1956). This paper was completed while the author held a fellowship at the Summer Research Institute of the Canadian Mathematical Congress, Kingston, 1957. The author is greatly indebted to Professors Dr. J. Popken (Amsterdam) and Dr. F. van der Blij (Utrecht) for valuable hints and kind assistance during the preparation of this paper.

which is equivalent to

$$\Omega Q_n(x) Q_m(x) = 0 \qquad n \neq m. \tag{1.4}$$

If R is the field of real numbers it is not true that each regular sequence $\{r_n\}$ can be obtained as a moment sequence of a non-negative distribution. Hence this notion of orthogonal polynomials is essentially more general than the usual one.

If the polynomial

$$Q_m(x) = \sum_{\mu=0}^{m} q_\mu x^\mu,$$

then from (1.3) it follows that the coefficients q_μ satisfy the equations (1.1); from the regularity of the sequence $\{r_n\}$ it follows that, apart from a constant factor, Q_m is uniquely determined.

If the orthogonal polynomials are considered in their monic form, that is, the coefficient of the highest power of x is one, then for the norm of Q_m, defined as $N_m = \Omega Q_m^2(x)$, we find that

$$N_m = q_0 r_m + q_1 r_{m+1} + \ldots + q_m r_{2m}, \qquad m \geqslant 0, \tag{1.5}$$

where again $Q_m(x) = \sum q_\mu x^\mu$.

These norms are not zero, for suppose $N_m = 0$, then the set (1.1) with $m + 1$ instead of m would have a non-trivial solution with $q_{m+1} = 0$.

2. Padé-approximants. Consider the formal Laurent series

$$D(x) = \sum_{\mu=-\infty}^{\infty} p_\mu x^\mu$$

with coefficients in R (in the sequel we will always suppress the limits of the summation index if these are $-\infty$ and $+\infty$). If n is any integer and m is a positive integer, the non-zero polynomial $V_{m,n}(x)$ of degree $\leqslant m$ will be called a *(Padé-) denominator of $D(x)$ for the place* (m, n), if in the formal product $V_{m,n}(x)\, D(x)$ the terms containing $x^{n+1}, x^{n+2}, \ldots, x^{n+m}$ have zero coefficients; any non-zero constant $V_{0,n}(x)$ will be called a denominator of $D(x)$ for the place $(0, n)$.

For each denominator $V_{m,n}(x)$ we define a *numerator* $U_{m,n}(x)$ as the series obtained from $V_{m,n}(x)\, D(x)$ by cancelling all terms after the one containing x^n. Then the pair $(U_{m,n}(x), V_{m,n}(x))$ will be called a *(Padé-)approximant* of $D(x)$ for the place (m, n).

We recall that if $D(x)$ is a formal *power* series over the real numbers, then $U_{m,n}(x)\, V_{m,n}(x)^{-1}$ is a Padé-fraction for the place (m, n) (cf., for example, Perron (**6**, §73)). Hence to any Padé-approximant of a real power series there corresponds a Padé-fraction.

If k is an arbitrary, but fixed, integer, a *sequence* of approximants belonging to the places $(m, m + k)$, $m = 0, 1, 2, \ldots$, will be called a *diagonal of order k*, denoted by D_k (the words "place" and "diagonal" are of course derived from the notion of the Padé-table (**6**)).

For each approximant $(U_{m,n}(x), V_{m,n}(x))$ we define an element $p_{m,n}$ (which may be zero) by means of the relation

(2.1) $\quad V_{m,n}(x)\, D(x) - U_{m,n}(x) = p_{m,n}x^{m+n+1} +$ higher powers of x.

We remark that for $m > 0$ the non-zero polynomial $V_{m,n}(x) = q_0x^m + \dots q_{m-1}x + q_m$ is a denominator of $\sum p_\mu x^\mu$ for the place (m, n) if and only if

$$(2.2) \qquad \begin{cases} q_0p_{n-m+1} + \dots + q_mp_{n+1} = 0 \\ \dots\dots\dots\dots\dots\dots \\ q_0p_n + \dots + q_mp_{n+m} = 0 \end{cases} .$$

Since the number of equations is always less than the number of unknowns, each place (m, n) has a Padé-approximant.

For the $p_{m,n}$ corresponding to this $V_{m,n}$ we find

$$(2.3) \qquad q_0p_{n+1} + \dots + q_mp_{n+m+1} = p_{m,n}.$$

If (U, V) and (U^*, V^*) are approximants for the same place, their sum, defined as $(U + U^*, V + V^*)$, is again an approximant for that place, and a constant multiple $p.(U, V)$, defined as (pU, pV), is an approximant for each place for which (U, V) is an approximant, except of course in the trivial case when this sum or multiple results in $(0, 0)$.

3. Regularity. The Padé-approximant $(U_{m,n}(x), V_{m,n}(x))$ of the series $D(x)$ for the place (m, n) will be called regular if (a) the constant term of $V_{m,n}(x)$ is 1, that is, $V_{m,n}(0) = 1$, and (b) any other approximant for the place (m, n) is a constant multiple of $(U_{m,n}, V_{m,n})$.

Clearly condition (b) is equivalent to the condition that any other denominator for the place (m, n) is a constant multiple of $V_{m,n}$.

The *denominator* of a regular approximant will also be called regular. A *set* of approximants will be called regular if each of its elements is regular.

For any place (m, n) there exists at most one regular approximant; 1 is a regular denominator for each place $(0, n)$ and for the corresponding $p_{0,n}$ we have $p_{0,n} = p_{n+1}$ if $D(x) = \sum p_\mu x^\mu$.

As an immediate consequence of our definition, we have

THEOREM 3.1. *The series $D(x) = \sum p_\mu x^\mu$ has a regular approximant for the place (m, n), $m > 0$, if and only if the set (2. 2), considered as equations in the q_μ, has exactly one solution with $q_m = 1$.*

If $(U_{m,n}, V_{m,n})$ is an approximant such that the corresponding $p_{m,n} = 0$, then $x(U_{m,n}, V_{m,n})$, defined as $(xU_{m,n}, xV_{m,n})$, is an approximant for the place $(m + 1, n + 1)$, the constant term of the denominator of which is 0; hence it cannot be a constant multiple of a regular approximant. This proves the following:

THEOREM 3.2. *If $D(x)$ has a regular diagonal D_k, then none of the corresponding $p_{m,m+k}$ is zero.*

If for a place (m, n) there exists no regular approximant, then there is an approximant for that place in which the denominator has constant term zero. For suppose that $V_{m,n}$ is a denominator such that $V_{m,n}(0) \neq 0$. Then there is certainly a $V^*_{m,n}$ that is no constant multiple of $V_{m,n}$. Then $V^*_{m,n}(x) - V^*_{m,n}(0)\ V_{m,n}(0)^{-1}\ V_{m,n}(x)$ is not zero, is a denominator for the place (m, n) and has constant term zero. This will be used in proving the following theorem, which is more or less a converse of Theorem 3.2:

THEOREM 3.3. *Let the series $D(x)$ have a diagonal D_k consisting of approximants $(U_{m,m+k}, V_{m,m+k})$ such that the corresponding $p_{m,m+k}$ are all different from zero, and $V_{m,m+k}(0) = 1$ for all m. Then D_k is regular.*

Proof. The approximant $(U_{0,k}, V_{0,k})$ has denominator 1, hence is regular. Suppose that for a certain m the approximant $(U_{m,m+k}, V_{m,m+k})$ is regular, and moreover that for the place $(m + 1, m + k + 1)$ there exists no regular approximant. Then there is an approximant (U^*, V^*) for this place such that $V^*(0) = 0$, hence $x^{-1}(U^*, V^*)$ is an approximant for the place $(m, m + k)$, and the corresponding $p^*_{m,m+k}$ is zero. However, since $p_{m,m+k} \neq 0$, this approximant cannot be a constant multiple of $(U_{m,m+k}, V_{m,m+k})$, contradicting the assumed regularity of the latter. Hence, if $(U_{m,m+k}, V_{m,m+k})$ is regular, there is a regular approximant for the place $(m + 1, m + k + 1)$, and $V_{m+1,m+k+1}$ is its denominator since $V_{m+1,m+k+1}(0) = 1$. Induction completes the proof.

We remark that the proof of this theorem shows that if a series $D(x)$ has a diagonal D_k for which the corresponding $p_{m,m+k}$ are all $\neq 0$, then all the approximants are constant multiples of regular approximants. Thus the condition $V_{m,m+k}(0) = 1$ in Theorem 3.3 has only a normalizing effect, and does not essentially restrict the class of series, which have a regular diagonal D_k, according to this theorem.

4. Approximants of $D(x^2)$. In this section we give a theorem relating approximants of the series $D(x^2)$ to those of the series $D(x)$.

THEOREM 4.1. *If the series $D(x) = \sum p_\mu x^\mu$ has regular diagonals D_k and D_{k+1} consisting of approximants $(U_{m,n}(x), V_{m,n}(x))$, then the series $D^*(x) = \sum p_\mu x^{2\mu}$ has a regular diagonal D_{2k+1} consisting of approximants $(U^*_{m,n}(x), V^*_{m,n}(x))$ for which we have*

$$(4.1)\qquad V^*_{2m,\,2m+2k+1}(x) = V_{m,m+k}(x^2), \qquad V^*_{2m+1,\,2m+2k+2}(x) = V_{m,m+k+1}(x^2),$$

together with analogous relations for U.

Moreover, if we put

$$V^*_{m,n}(x)\, D^*(x) - U^*_{m,n}(x) = p^*_{m,n}\, x^{m+n+1} + \dots$$

then

$$(4.2)\qquad p^*_{2m,\,2m+2k+1} = p_{m,m+k}, \qquad p^*_{2m+1,\,2m+2k+2} = p_{m,m+k+1}.$$

https://doi.org/10.4153/CJM-1958-060-2 Published online by Cambridge University Press

Proof. Consider any approximant $(U(x), V(x))$ of $D(x^2)$ for the place $(2m+1, 2m+2k+2)$; then

$$V(x)D(x^2) - U(x) = px^{4m+2k+4} + qx^{4m+2k+5} + \dots .$$

Now $V(x)$ can be written as $V_m'(x^2) + x\, V_m''(x^2)$, where the subscript denotes the highest power of the polynomial variable. Similarly

$$U(x) = U'_{m+k+1}(x^2) + x\, U''_{m+k}(x^2).$$

Then we have

$$V'_m(x^2)D(x^2) - U'_{m+k+1}(x^2) = p\, x^{4m+2k+4} + \dots ,$$
$$xV''_m(x^2)D(x^2) - xU''_{m+k}(x^2) = q\, x^{4m+2k+5} + \dots ,$$

or

$$V'_m(x)\, D(x) - U'_{m+k+1}(x) = p\, x^{2m+k+2} + \dots , \tag{4.3}$$
$$V''_m(x)\, D(x) - U''_{m+k}(x) = q\, x^{2m+k+2} + \dots . \tag{4.4}$$

From the regularity of D_{k+1} it follows that, apart from an arbitrary constant factor, there is exactly one $V_m'(x)$ satisfying (4.3), viz. $V_{m,m+k+1}(x)$ and that $p = p_{m,m+k+1}$ if $V_m' = V_{m,m+k+1}$.

On the other hand, from Theorem 3.2 and the regularity of D_k it follows that the only $V_m''(x)$ satisfying (4.4) is the zero polynomial. Hence all possible $V(x)$ are constant multiples of $V_{m,m+k+1}(x^2)$. Since, moreover, $V_{m,m+k+1}(0) = 1$, it follows that $V^*_{2m+1,2m+2k,2}(x) = V_{m,m+k+1}(x^2)$ is regular. It follows also that $p^*_{2m+1,2m+2k+2} = p_{m,m+k+1}$. The remaining part of the theorem can be proved in a similar way.

It may be shown that no series $D(x^n)$, with $n > 2$, can have a regular diagonal.

5. Recurrence relations for approximants.

THEOREM 5.1. *Let the series $D(x)$ have diagonals D_k and D_{k+1} consisting of approximants $(U_{m,n}, V_{m,n})$ such that D_k is regular and $V_{m,m+k+1}(0) = 1$ for all m. Then we have*

$$U_{m+1,m+k+1} = U_{m,m+k+1} - p_{m,m+k+1}\, p^{-1}_{m,m+k} U_{m,m+k} \tag{5.1}$$
$$V_{m+1,m+k+1} = V_{m,m+k+1} - p_{m,m+k+1}\, p^{-1}_{m,m+k} V_{m,m+k}. \tag{5.2}$$

Proof. The highest powers in the right-hand members of (5.1) and (5.2) are at most $m+k+1$ and $m+1$ respectively, whereas the right-hand side of (5.2) cannot be identically zero since it has constant term 1. Finally, multiplying the right-hand side of (5.2) by $D(x)$ and subtracting the right-hand side of (5.1) shows that the right-hand members of (5.1) and (5.2) constitute an approximant for the place $(m+1, m+k+1)$. Since D_k is regular the relations (5.1) and (5.2) must be true.

In a similar way one can prove

https://doi.org/10.4153/CJM-1958-060-2 Published online by Cambridge University Press

THEOREM 5.2. *Let the series $D(x)$ have diagonals D_k and D_{k+1} consisting of approximants $(U_{m,n}, V_{m,n})$ such that D_{k+1} is regular and $V_{m,m+k}(0) = 1$ for all m. Then we have*

$$V_{m+1,m+k+2} = V_{m+1,m-k+1} - p_{m+1,m+k+1}\, p_{m,m+k+1}^{-1} V_{m,m+k+1}, \tag{5.3}$$

together with an analogous relation for U.

Combining these two theorems we obtain

THEOREM 5.3. *Let the series $D(x)$ have regular diagonals D_k and D_{k+1} consisting of approximants $(U_{m,n}, V_{m,n})$. Then for all m we have*

$$V_{m+1,m+k+1} = (1 - a_m x) V_{m,m+k} - b_m x^2\, V_{m-1,m+k-1} \tag{5.4}$$

where

$$a_m = p_{m,m+k}\, p_{m-1,m+k}^{-1} + p_{m,m+k+1}\, p_{m,m+k}^{-1} \tag{5.5}$$

$$b_m = p_{m,m+k}\, p_{m-1,m+k-1}^{-1}. \tag{5.6}$$

An analogous relation for U holds likewise.

6. Orthogonality relations of approximants. Let $D(x) = \sum p_\mu x^\mu$ have a regular diagonal D_k consisting of approximants $(U_{m,m+k}, V_{m,m+k})$. If we write $V_{m,m+k} = q_0 x^m + q_1 x^{m-1} + \ldots + q_m$ where the q_μ of course depend on m, then the coefficients satisfy the equation (2.2) with $n = m + k$. By the substitution

$$p_{k+1} = r_0, \quad p_{k+2} = r_1, \quad p_{k+3} = r_2, \ldots \tag{6.1}$$

this system is transformed into (1.1), the condition for orthogonality. Since D_k is regular, it follows that for each $m > 0$ and $n = m + k$ the set (2.2) has exactly one solution with $q_m = 1$, hence the same is true for the related system (1.1). Hence the polynomials $q_0 + q_1 x + \ldots + q_m x^m = x^m V_{m,m+k}(x^{-1})$ form a set of orthogonal polynomials with respect to the sequence p_{k+1}, p_{k+2}, $p_{k+3}, \ldots$, and are monic by virtue of $V_{m,m+k}(0) = 1$.

The same substitution (6.1) transforms (2.3) into (1.5). Hence the norm of $x^m V_{m,m+k}(x^{-1})$ is $p_{m,m+k}$. This proves

THEOREM 6.1. *Let the series $D(x) = \sum p_\mu x^\mu$ have a regular diagonal D_k consisting of approximants $(U_{m,m+k}, V_{m,m+k})$; then the polynomials $V_m(x)$ defined by*

$$V_m(x) = x^m V_{m,m+k}(x^{-1}), \qquad m = 0, 1, 2, \ldots \tag{6.2}$$

are monic, form an orthogonal set with respect to the sequence $p_{k+1}, p_{k+2}, p_{k+3}, \ldots$, and have norms $p_{m,m+k}$.

Since the sequence $V_{m,m+k}$ forms a diagonal, the $V_m(x)$ will be called *diagonal polynomials.*

If the sequence $D(x) = \sum p_\mu x^\mu$ has regular diagonals D_k and D_{k+1}, then from Theorem 4.1 it follows that $D^*(x) = \sum p_\mu x^{2\mu}$ has a regular diagonal

D^*_{2k+1}, and from Theorem 6.1 and the relations (4.1) and (4.2) we then have the following:

THEOREM 6.2. *Let the series* $D(x) = \sum p_\mu x^\mu$ *have regular diagonals* D_k *and* D_{k+1} *consisting of approximants* $(U_{m,n}(x), V_{m,n}(x))$; *then the polynomials* $W_m(x)$ *defined by*

$$W_{2m}(x) = x^{2m} V_{m,m+k}(x^{-2}), \; W_{2m+1}(x) = x^{2m+1} V_{m,m+k+1}(x^{-2}) \tag{6.3}$$

are monic, form an orthogonal set with respect to the sequence p_{k+1}, 0, p_{k+2}, 0, p_{k+3}, 0, . . . , *and have norms* $\Omega(W_{2m}{}^2) = p_{m,m+k}$, $\Omega(W_{2m+1}{}^2) = p_{m,m+k+1}$.

These polynomials will be called *stepline polynomials* (since the sequence $V_{0,k}$, $V_{0,k+1}$, $V_{1,k}$, $V_{1,k+1}$, . . . , forms a stepline in the Padé-table).

Combination of Theorems 6.1 and 5.3 gives

THEOREM 6.3. *Under the conditions of Theorem* 6.2 *the assertions of Theorem* 6.1 *hold, and the polynomials* $V_m(x)$ *satisfy the recurrence relation*:

$$V_{m+1}(x) = (x - a_m)\, V_m(x) - b_m\, V_{m-1}(x), \qquad m \geqslant 1, \tag{6.4}$$

where a_m *and* b_m *are given by* (5.5) *and* (5.6).

Combination of Theorems 6.2, 5.1, and 5.2 gives

THEOREM 6.4. *Under the conditions of Theorem* 6.2 *the assertions of Theorem* 6.2 *hold, and the polynomials* $W_m(x)$ *satisfy the recurrence relations*

$$W_{2m+1} = x\, W_{2m} - p_{m,m+k}\, p^{-1}_{m-1,m+k}\, W_{2m-1}, \qquad m \geqslant 1, \tag{6.5}$$

$$W_{2m+2} = x\, W_{2m+1} - p_{m,m+k+1}\, p^{-1}_{m,m+k}\, W_{2m}, \qquad m \geqslant 0. \tag{6.6}$$

7. Extension to rings. Hitherto it has been assumed that we are working in a commutative field. It may now be pointed out that a similar theory exists if R is a not necessarily commutative ring with unit element. This is of importance when we consider matrix polynomials, that is, polynomials having matrix-coefficients. But in this case the orthogonality according to (1.3) is not complete since (1.4) is only true for $n < m$. However, if we take the matrices r_ν hermitian and for any two polynomials $P(x) = \sum p_\mu x^\mu$, $Q(x) = \sum q_\mu x^\mu$ define an "inner product" $\{P, Q\} = \sum p_\mu r_{\mu+\nu} q^*_\nu$ then for any $m \neq n$ we have $\{Q_m, Q_n\} = 0$ instead of (1.4). This is quite natural in connection with complex valued orthogonal functions, where $f(x)$ and $g(x)$ are called orthogonal with respect to a real non-negative distribution $d\psi(x)$ if $\int f(x) d\psi(x)\, \bar{g}(x) = 0$.

Returning to the general case where R is an arbitrary ring, we see that all notions in §§1-6 have a meaning. To render this true, we accept such rules as: the set of equations (1.1) has exactly one solution with $q_m = 1$. However it is no longer true that to each place there corresponds an approximant. If we call a quantity b a right zero divisor if there is a $c \neq 0$ such that $cb = 0$, then the norms of a set of orthogonal polynomials are not even right zero divisors. In the definition of regularity (§3) a constant left-multiple is required. Theorem 3.1 remains true; in Theorem 3.2 the $p_{m,m+k}$ are not right zero divisors, and Theorem 3.3 remains true under the additional condition that the $p_{m,m+k}$ should not be right zero divisors. The final remark of §3 is no longer true. It may be remarked that the conditions of Theorem 3.3, with $p_{m,m+k}$ not right zero divisors, are much weaker than those in the corresponding Theorem 6.3 of the author's thesis.

All the other theorems in Part I remain true, provided only that such expressions as $p_{m,m+k}\, p_{m,m+k-1}{}^{-1}$, if present, exist; here ab^{-1} will be said to exist and be equal to c if b is not a right zero divisor and $a = cb$.

PART II

8. General hypergeometric series. Let us compare the definition of the ordinary hypergeometric series

$$(8.1)\quad F(a, b; c; x) = \sum_{\mu=0}^{\infty} \frac{a(a+1)\ldots(a+\mu-1)b\ldots(b+\mu-1)}{c(c+1)\ldots(c+\mu-1)\mu!}x^{\mu}$$

with that of the Heine series

$$(8.2)\quad H(a, b; c; x) =$$

$$\sum_{\mu=0}^{\infty} \frac{(1-a)(1-aq)\ldots(1-aq^{\mu-1})(1-b)\ldots(1-bq^{\mu-1})}{(1-c)(1-cq)\ldots(1-cq^{\mu-1})(1-q)\ldots(1-q^{\mu})}x^{\mu},$$

where q is not a root of unity.

We remark that in both cases the coefficient of x^n can be represented as

$$(8.3)\quad \frac{[a, 0][a, 1]\ldots[a, n-1][b, 0][b, 1]\ldots[b, n-1]}{[c, 0][c, 1]\ldots[c, n-1][\sigma, 1][\sigma, 2]\ldots[\sigma, n]}$$

where

$$(8.4)\quad [s, k] = s + k, \qquad \sigma = 0,$$

or

$$(8.5)\quad [s, k] = 1 - sq^k, \qquad \sigma = 1.$$

In both cases we have

$$(8.6)\quad [\sigma, 0] = 0$$

$$(8.7)\quad [\sigma, h] \neq 0 \text{ for all } h \neq 0$$

and

$$(8.8)\quad [s, k+h] = g_1(h)[s, k] + g_2(h)$$

for all h, k and s, where $g_1(h)$ and $g_2(h)$ are functions of h only.

To obtain a unified treatment of hypergeometric and Heine series we shall assume that $[s, k]$ is any function with values in a commutative field R which is defined for all elements s of a set S (that need not consist of elements of R) and all integers k. Assume also that $[s, k]$ satisfies (8.6) and (8.7) for a certain element $\sigma \in S$, and satisfies (8.8) for all $s \in S$ and all integers h and k.

By (8.6) and (8.7), the condition (8.8) is equivalent to

$$(8.9)\qquad \begin{vmatrix} 1 & [a,\ k] & [a,\ k+h] \\ 1 & [b,\ 1] & [b,\ \ l+h] \\ 1 & [c, m] & [c, m+h] \end{vmatrix} = 0 \text{ for all } a,\ b,\ c \in S \text{ and all } h,\ k,\ l,\ m.$$

In fact, if (8.8) holds then the columns in (8.9) are linearly dependent. Conversely, if in (8.9) we substitute $a = s$, $b = c = \sigma$, $m = 0$, $l = -h \neq 0$ we obtain

$$(8.10)\qquad [s, k+h] = -[\sigma, h][\sigma, -h]^{-1}[s, k] + [\sigma, h],$$

which clearly has the form (8.8). It follows that for $h \neq 0$ we have

$$g_1(h) = -[\sigma, h][\sigma, -h]^{-1},\ g_2(h) = [\sigma, h],$$

whereas $g_1(0) = 1$, $g_2(0) = 0$.

To simplify our notation, we define

$$(8.11)\qquad [s, k]_0 = 1$$

for all k and s,

$$(8.12)\qquad [s, k]_h = [s, k][s, k+1] \ldots [s, k+h-1]$$

for all $h > 0$, k and s

$$(8.13)\qquad [s, k]_{-h} = [s, k-1]^{-1} \ldots [s, k-h]^{-1}$$

for all $h > 0$, k and s for which the right-hand member is defined.

Then we have

$$(8.14)\qquad [s, k]_h = [s, k]_m\, [s, k+m]_{h-m}$$

for all h, k, m and s for which the right-hand member is defined.

The series

$$(8.15)\qquad F([a, k], [b, l]; [c, m]; x) = \sum_{\mu=0}^{\infty} \frac{[a, k]_\mu [b, l]_\mu}{[c, m]_\mu [\sigma, l]_\mu} x^\mu,$$

which is defined if $[c, m+\mu] \neq 0$ for all $\mu \geqslant 0$, will be called the general hypergeometric series.

If $[s, k]$ and σ are given by (8.4) or (8.5), the general hypergeometric series is an ordinary hypergeometric or Heine series, respectively. Since $[c, m+\mu] \neq 0$ for all $\mu \geqslant 0$, we have the formal identity

$$(8.16)\qquad F([a, k], [b, l]; [c, m]; x) - F([a, k+1], [b, l]; [c, m+1]; x) + [c, m]_2^{-1}[b, l]\,\{[c, m] - [a, k]\}F([a, k+1], [b, l+1]; [c, m+2]; x) = 0.$$

For the constant terms cancel, whereas the coefficient of x^μ, $\mu > 0$, is equal to

$$[c, m]_{\mu+1}^{-1}\,[\sigma, 1]_\mu^{-1}\,[a, k+1]_{\mu-1}\,[b, l]_\mu \\ \{[a, k][c, m+\mu] - [a, k+\mu][c, m] + ([c\ m] - [a, k])[\sigma, \mu]\},$$

which is zero by virtue of (8.10).

https://doi.org/10.4153/CJM-1958-060-2 Published online by Cambridge University Press

In a similar way it can be verified that

(8.17) $F([a, k+1], [b, l]; [c, m]; x) - F([a, k], [b, l+1]; [c, m]; x) + [c, m]^{-1}\{[a, k] - [b, l]\}F([a, k+1], [b, l+1]; [c, m+1]; x) = 0$

if $[c, m+\mu] \neq 0$ for all $\mu \geqslant 0$.

9. The functions $[s, k]$**.** In this section we investigate the functions $[s, k]$ satisfying equations (8.6) to (8.8).

If we put

(9.1) $$-[\sigma, 1][\sigma, -1]^{-1} = q$$

then from (8.10) for $h = 1$, $s = \sigma$, it follows that

(9.2) $$[\sigma, k+1] = q[\sigma, k] + [\sigma, 1].$$

Hence by induction

(9.3) $$[\sigma, k+h] = q^h[\sigma, k] + (q^{h-1} + q^{h-2} + \ldots + q + 1)[\sigma, 1] \quad \text{for } h \geqslant 0.$$

Putting $k = 0$ in (9.3) we obtain

(9.4) $$[\sigma, h] = (q^{h-1} + q^{h-2} + \ldots q + 1)[\sigma, 1] \qquad \text{for } h \geqslant 0;$$

putting $k = -h$ in (9.3) we obtain

(9.5) $$[\sigma, -h] = -q^{-h}(q^{h-1} + q^{h-2} + \ldots + q + 1)[\sigma, 1] \qquad \text{for } h \geqslant 0.$$

Conversely, the relations (9.4) and (9.5) define the function $[\sigma, h]$ if $[\sigma, 1]$ and q are given. These elements can be chosen quite arbitrarily. We formulate this as a theorem:

THEOREM 9.1. *If* $[\sigma, 1]$ *and* $q \neq 0$ *are given elements, then the function* $[\sigma, h]$ *defined by* (9.4) *and* (9.5) *satisfies* (8.6) *and* (8.8) *for* $s = \sigma$, *and all* h, k. *In order that this function satisfy* (8.7) *also,* $[\sigma, 1]$ *should be chosen* $\neq 0$ *and* q *should not be a root of unity different from* 1.

Another consequence of (9.4) and (9.5) is that

(9.6) $$[\sigma, -h] = -q^{-h}[\sigma, h] \qquad \text{for all } h.$$

Substituting this result in (8.10) we obtain

(9.7) $$[s, k+h] = q^h[s, k] + [\sigma, h].$$

Hence

(9.8) $$[a, k+h] - [c, l+h] = q^h\{[a, k] - [c, l]\}.$$

From (9.7) we have the special case

(9.9) $$[s, h] = q^h[s, 0] + [\sigma, h],$$

hence the function $[s, h]$ is determined as soon as the functions $[s, 0]$ and $[\sigma, h]$ are known (q being defined by (9.1)). This is formulated in the following theorem, which is readily verified:

THEOREM 9.2. *If* $[\sigma, h]$ *is any function satisfying conditions* (8.6) to (8.8) *for* $s = \sigma$, *and all* h, k, *and* $[s, 0]$ *is any function defined for all* s *in some set* S (*of course* $[\sigma, 0] = 0$), *then the function* $[s, h]$ *defined by* (9.9) *satisfies* (8.8) (*and of course* (8.6) *and* (8.7)) *for all* s, h, k.

This theorem enables us to extend the domain S of a function $[s, h]$ satisfying equations (8.6) to (8.8): suppose we add an element a to S and take for $[a, 0]$ an arbitrary element of R. Then, if we put $[a, k] = q^k[a, 0] + [\sigma, k]$ in accordance with (9.9), it follows that the extended function $[s, k]$ still satisfies equations (8.6) to (8.8).

We shall apply this in particular by first extending the field R, mentioned in the definition of the function $[s, k]$ (cf. §8), to the field $R(y)$ of rational functions in the variable y over R. Then we add the element y to S and define $[y, 0] = y$, hence

$$[y, k] = q^k y + [\sigma, k] \tag{9.10}$$

(cf. §10). From (9.7) with (c, m, k) instead of (s, k, h) it follows that, if in the right-hand member of (9.10) we substitute $[c, m]$ for $y = [y, 0]$, then $[y, k]$ is replaced by $[c, k + m]$. Hence, if in $[y, 0]_\mu$, considered as a polynomial in y, we substitute $[c, m]$ for y, this expression is transformed into $[c, m]_\mu$.

In later sections we will add y to S, but take $[y, 0] = y^{-1}$, hence

$$[y, k] = q^k y^{-1} + [\sigma, k]. \tag{9.11}$$

From Theorems 9.1 and 9.2 it follows that the class of all possible functions $[s, h]$ is very limited. Actually (9.4) and (9.5) simplify to

$$[\sigma, k] = [\sigma, 1] \,.\, k \qquad \text{if } q = 1, \tag{9.12}$$

$$[\sigma, k] = [\sigma, 1] \frac{1 - q^k}{1 - q} \qquad \text{if } q \neq 1. \tag{9.13}$$

Hence:

THEOREM 9.3. *The function* $[s, k]$ *satisfies*

$$[s, k] = [s, 0] + [\sigma, 1] \,.\, k \qquad \text{if } q = 1, \tag{9.14}$$

or

$$[s, k] = q^k[s, 0] + [\sigma, 1] \frac{1 - q^k}{1 - q} \qquad \text{if } q \neq 1 \tag{9.15}$$

(*in the latter case* q *cannot be a root of unity*).

From this theorem it follows that if $q = 1$, the series (8.15) always is an ordinary hypergeometrical series, whereas if $q \neq 1$ the series is a Heine series. However, as we wish to treat the hypergeometric and Heine series simultaneously, we shall not make use of this theorem in the present part. It has, however, some importance in connection with Part III.

10. Hypergeometric polynomials. Since $[\sigma, 0] = 0$ it follows that $[\sigma, -n]_\mu = 0$ for $\mu > n \geqslant 0$. Hence $F([\sigma, -n], [b, l]; [c, m]; x)$ is a polynomial of degree $\leqslant n$. These polynomials will be called general hypergeometric polynomials.

In the following we shall consider expressions such as

$$[c, m]_n\, F([\sigma, -n], [b, l]; [c, m]; x), \qquad n \geqslant 0. \tag{10.1}$$

If $[c, m + \mu]$ has not an inverse for all $\mu \geqslant 0$, we can still give a meaning to (10.1) by extending R to $R(y)$, defining $[y, k]$ by (9.10), and substituting $[c, m]$ for y in

$$[y, 0]_n\, F([\sigma, -n], [b, l]; [y, 0]; x).$$

This definition is consistent with the usual meaning of (10.1) if $[c, m + \mu]$ has an inverse for all $\mu \geqslant 0$.

For polynomials of type (10.1) we have the following generalization of (8.17):

$$\begin{aligned} &[c, m]_n\, F([\sigma, -n + 1], [b, l]; [c, m]; x) \\ &- [c, m]_n\, F([\sigma, -n], [b, l + 1]; [c, m]; x) \\ &+ ([\sigma, -n] - [b, l])\, [c, m + 1]_{n-1} F([\sigma, -n + 1], [b, l + 1]; [c, m + 1]; x) \\ &= 0, \end{aligned} \tag{10.2}$$

which holds for all b and c, $n > 0$, m and l.

Let S' be the subset of S consisting of those elements s, to each of which there corresponds a uniquely determined element $s' \in S$ such that for each k

$$[s', -k] = f(s)\, q^{-k}[s, k], \tag{10.3}$$

where $f(s)$ does not depend on k.

Then it is only a matter of calculation to show that, if $a, c \in S'$, $m \geqslant 0$ and the left-hand member exists,

$$\begin{aligned} x^m F([\sigma, -m], [a, k]; [c, l]; x^{-1}) &= [a, k]_m [c, l]_m^{-1} q^{-\frac{1}{2}m(m+1)}(-)^m \\ &\times F([\sigma, -m], [c', -l - m + 1]; [a', -k - m + 1]; x f(a) f(c)^{-1} q^{m+l-k+1}). \end{aligned} \tag{10.4}$$

The set S' is not empty, since (10.3) holds for $s = s' = \sigma$, $f(\sigma) = -1$. Actually, it is easy to see that under some circumstances S' may contain more than one element of S, or may coincide with S.

11. Padé's theorem. In this section we shall formulate and prove a generalization of a result of Padé **(5)**. In the following, j denotes an arbitrary but fixed integer which may be $-\infty$.

THEOREM 11.1. *Let the series*

$$F(x) = \sum_{\mu=j}^{\infty} \frac{[a, 0]_\mu}{[c, 0]_\mu} x^\mu \tag{11.1}$$

be defined. If for $0 \leqslant m \leqslant n - j + 1$ *we put*

$$\begin{aligned} V_{m,n}(x) &= (-1)^m \frac{[a, n - m + 1]_m}{[c, n]_m} q^{\frac{1}{2}m(m-1)} x^m \\ &\times F([\sigma, -m], [c, n]; [a, n - m + 1]; x^{-1} q), \end{aligned} \tag{11.2}$$

https://doi.org/10.4153/CJM-1958-060-2 Published online by Cambridge University Press

then $V_{m,n}(x)$ *is a Padé denominator of* $F(x)$ *for the place* (m, n). *Moreover, if we put*

$$(11.3) \quad p_{m,n} = \frac{[\sigma, 1]_m [a, 0]_{n+1}}{[c, 0]_{m+n+1}[c, n]_m} q^{mn}\{[c, 0] - [a, 0]\} \ldots \{[c, m-1] - [a, 0]\}$$

and

$$(11.4) \qquad F_{m,n}(x) = F([a, n+1], [\sigma, m+1]; [c, m+n+1]; x),$$

then there exists a series $U_{m,n}(x)$ *in which the highest exponent of* x *is* $\leqslant n$, *such that*

$$(11.5_{m,n}) \qquad V_{m,n}(x)\, F(x) = U_{m,n}(x) + p_{m,n} x^{m+n+1} F_{m,n}(x).$$

Proof. The existence of the series (11.1) implies the existence of $\{[c, n]_m\}^{-1}$ in (11.2), and then from §10 it follows that $V_{m,n}(x)$ is defined and represents a polynomial of degree $\leqslant m$. It also follows from the existence of (11.1) that $p_{m,n}$ and $F_{m,n}$ are defined. Then $(11.5_{m,n})$ implies that $V_{m,n}(x)$ is a Padé-denominator of $F(x)$ for the place (m, n). The proof of $(11.5_{m,n})$ will be performed by induction.

Firstly, $(11.5_{0,n})$ is true for all $n \geqslant j - 1$.

From (10.2) with $(c, a, m, n-1, n-m+1, x^{-1}q)$ instead of (b, c, n, l, m, x) it follows after some calculations that

$$(11.6_{m,n}) \qquad V_{m,n}(x) = V_{m-1,n}(x) - q_{m,n} x\, V_{m-1,n-1}(x).$$

where $q_{m,n} = [c, n+m-2]_2^{-1}\, [a, n][c, n-1]q^{m-1}$, and these $q_{m,n}$ satisfy

$$(11.7_{m,n}) \qquad p_{m-1,n-1}\, q_{m,n} = p_{m-1,n}.$$

From (8.16) it follows after some calculations that

$$(11.8_{m,n}) \qquad F_{m-1,n-1}(x) = F_{m-1,n}(x) - r_{m,n} F_{m,n}(x)$$

where, by (9.8),

$$r_{m,n} = [c, m+n-1]_2^{-1}\, [\sigma, m]\, \{[c, m-1] - [a, 0]\}\, q^n,$$

and these $r_{m,n}$ satisfy

$$(11.9_{m,n}) \qquad p_{m-1,n}\, r_{m,n} = p_{m,n}.$$

Now suppose that for certain integers m and n the relations $(11.5_{m-1,n-1})$ and $(11.5_{m-1,n})$ have already been proved. Then using successively $(11.6_{m,n})$, $(11.5_{m-1,n})$, $(11.5_{m-1,n-1})$, $(11.7_{m,n})$, $(11.8_{m,n})$ and $(11.9_{m,n})$, and putting $U_{m,n}(x) = U_{m-1,n}(x) - q_{m,n}x\, U_{m-1,n-1}(x)$ we obtain

$$\begin{aligned} V_{m,n}(x)\, F(x) &= V_{m-1,n}(x)\, F(x) - q_{m,n}x\, V_{m-1,n-1}(x)\, F(x) \\ &= U_{m-1,n}(x) + p_{m-1,n}x^{m+n}\, F_{m-1,n}(x) - q_{m,n}x\, U_{m-1,n-1}(x) \\ &\qquad - q_{m,n}p_{m-1,n-1}x^{m+n}\, F_{m-1,n-1}(x) \\ &= U_{m-1,n}(x) - q_{m,n}x\, U_{m-1,n-1}(x) \\ &\qquad + p_{m-1,n}x^{m+n}\, (F_{m-1,n}(x) - F_{m-1,n-1}(x)) \\ &= U_{m,n}(x) + p_{m-1,n}\, r_{m,n}\, x^{m+n+1}\, F_{m,n}(x) \\ &= U_{m,n}(x) + p_{m,n}\, x^{m+n+1}\, F_{m,n}(x). \end{aligned}$$

https://doi.org/10.4153/CJM-1958-060-2 Published online by Cambridge University Press

Thus $(11.5_{m,n})$ follows from $(11.5_{m-1,n-1})$ and $(11.5_{m-1,n})$. Since $(11.5_{0,n})$ is satisfied for all $n \geqslant j-1$, it follows that $(11.5_{m,n})$ is true for $0 \leqslant m \leqslant n-j+1$.

Remarks. I. From (11.2) it follows that $V_{m,n}(0) = 1$.
II. Padé's original theorem (cf. (5)) gives denominators for a certain class of hypergeometric series. This corresponds to the special case $j = 0$ of our theorem: we then have $F(x) = F([a, 0], [\sigma, 1]; [c, 0]; x)$. For every finite j we have

$$(11.10) \qquad F(x) = [a, 0]_j [c, 0]_j^{-1} x^j F([a, j], [\sigma, 1]; [c, j]; x).$$

III. If $a, c \in S'$ (cf. §10), then from (11.2) and (10.5) it follows that

$$(11.11) \quad V_{m,n}(x) = F([\sigma, -m], [a', -n]; [c', -m-n+1]; xqf(c)f(a)^{-1}).$$

In particular we see that the ordinary hypergeometric polynomial $F(-m, -a-n; -c-m-n+1; x)$ is a Padé-denominator of the ordinary hypergeometric series $F(a, 1; c; x)$ for the place (m, n) if $m \leqslant n+1$. This is Padé's original theorem. We also see that the Heine polynomial (for notation cf. (8.2)) $H(q^{-m}, a^{-1}q^{-n}; c^{-1}q^{-m-n+1}; xac^{-1}q)$ is a Padé-denominator of the Heine series $H(a, q; c; x)$ for the place (m, n) if $m \leqslant n+1$.

12. Confluent series. Consider the rational function field $R(y)$, and, as in (9.11), let $[y, 0] = y^{-1}$ and $[y, k] = q^k y^{-1} + [\sigma, k]$. This expression has an inverse in $R(y)$ for every k. Then the coefficient of x^μ in the series $F([a, k], [b, l]; [y, 0]; xy^{-1})$ is

$$[a, k]_\mu [b, l]_\mu \{(q + [\sigma, 1]y) \ldots (q^{\mu-1} + [\sigma, \mu-1]y)[\sigma, 1]_\mu\}^{-1}.$$

The substitution $y = 0$ in this expression gives a meaningful result. By this substitution we obtain from the given series the series

$$_2F_0([a, k], [b, l]; xl) = \sum_{\mu=0}^{\infty} \frac{[a, k]_\mu [b, l]_\mu}{[\sigma, 1]_\mu} q^{-\frac{1}{2}\mu(\mu-1)} x^\mu.$$

In a similar way we obtain by putting $y = 0$ in $F([a, k], [y, 0]; [c, m]; xy)$ the series

$$_1F_1([a, k]; [c, m]; x) = \sum_{\mu=0}^{\infty} \frac{[a, k]_\mu}{[c, m]_\mu [\sigma, 1]_\mu} q^{\frac{1}{2}\mu(\mu-1)} x^\mu.$$

The series $_2F_0$ and $_1F_1$ will be called confluent series of the first kind. It is also easily verified that $y = 0$ in $F([a, k], [b, l]; [y, m]; xy^{-1})$ gives $_2F_0([a, k], [b, l]; xq^{-m})$, and that $y = 0$ in $F([a, k], [y, l]; [c, m]; xy)$ gives $_1F_1([a, k]; [c, m]; xq^1)$.

If $q \neq 1$ then for a special value of one of the parameters in $F([a, k], [b, 1]; [b, l]; [c, m]; x)$ we get series which are intimately connected with these confluent series of the first kind. In fact let $[\zeta, 0] = [\sigma, 1]\ (1-q)^{-1}$. Then $[\zeta, \mu] = [\zeta, 0]$ for all μ; hence

$$F([a, k], [b, l]; [\zeta, m]; x[\zeta, 0]) = \sum_{\mu=0}^{\infty} \frac{[a, k]_\mu [b, l]_\mu}{[\sigma, 1]_\mu} x^\mu.$$

This leads us to the definition of the confluent series of the second kind:

$$
\begin{aligned}
{}_2F_0^*([a, k], [b, l]; x) &= F([a, k], [b, l]; [\zeta, 0]; x[\zeta, 0]),\\
{}_1F_1^*([a, k]; [c, m]; x) &= F([a, k], [\zeta, 0]; [c, m]; x[\zeta, 0]^{-1}),\\
{}_1F_0^*([a, k]; x) &= {}_1F_1([a, k]; [\zeta, 0]; x[\zeta, 0]) = {}_2F_0^*([a, k], [y, 0]; xy)_{y=0},\\
{}_1F_0^{**}([a, k]; x) &= {}_2F_0([a, k], [\zeta, 0]; x[\zeta, 0]^{-1}) = {}_1F_1^*([a, k]; [y, 0]; xy^{-1})_{y=0}.
\end{aligned}
$$

The initial assumption $q \neq 1$ can be avoided by considering q as a variable over R, in which case the defining expressions above for the four types of confluent series of the second kind certainly have a sense, and allow substitution of any value for q which satisfies the conditions in Theorem 9.1. If $q = 1$ we again obtain the former confluent series.

The close relationship between the series of different kinds arises from (10.4). In fact, if we apply confluence of any kind to (10.4), then on the left- and right-hand sides confluent series of opposite kinds appear.

From Theorem 11.1 it is easy to deduce corresponding theorems for the confluent series. If we replace x by xy^{-1}, take $c = y$, where $[y, \mu] = q^\mu y^{-1} + [\sigma, k]$ and put $y = 0$, we obtain the following:

THEOREM 12.1. *Let the series $F'(x)$ be defined by*

$$F'(x) = \sum_{\mu=j}^{\infty} [a, 0]_\mu q^{-\frac{1}{2}\mu(\mu-1)} x^\mu$$

for some integer j. If for $0 \leqslant m \leqslant n - j + 1$ we put

$$
\begin{aligned}
V'_{m,n}(x) &= (-1)^m [a, n - m + 1]_m q^{-mn} x^m {}_1F_1([\sigma, -m]; [a, n - m + 1]; x^{-1} q^{n+1})\\
p'_{m,n} &= [\sigma, 1]_m [a, 0]_{n+1} q^{-\frac{1}{2}(m+n)(m+n+1)}\\
F'_{m,n}(x) &= {}_2F_0([a, n + 1], [\sigma, m + 1]; xq^{-m-n-1})
\end{aligned}
$$

then there exists a series $U_{m,n}'(x)$ of degree $\leqslant n$ such that

$$V'_{m,n}(x) F'(x) = U'_{m,n}(x) + p'_{m,n} x^{m+n+1} F'_{m,n}(x).$$

Similar results can be obtained for all the other cases.

13. Generalized classical orthogonal polynomials. I. If the general hypergeometric series has regular diagonals, then from Theorem 11.1 and Theorems 6.1 and 6.2 we can derive explicit forms for the corresponding orthogonal polynomials.

Now, for all m, n such that $0 \leqslant m \leqslant n - j + 1$, the $V_{m,n}$ given by (11.2) have $V_{m,n}(0) = 1$, and hence, by virtue of Theorem 3.3 those $V_{m,m+k}$ will constitute a regular diagonal if the corresponding $p_{m,m+k}$ are different from zero; by virtue of Theorem 3.2 this condition is also necessary, and then from (11.3) it follows that $[a, \mu]$ and $\{[c, \mu] - [a, 0]\}$ should be different from zero for all $\mu \geqslant 0$. However, this condition is independent of k, and hence all diagonals of order k, $k \geqslant j - 1$, are regular at the same time. Hence Theorems 6.3 and 6.4 are applicable if this condition is satisfied. From Theorem 6.3 we obtain

THEOREM 13.1. *Let k be an arbitrary integer; let $[a, \mu]$, $[c, \mu]$ and $\{[c, \mu] - [a, 0]\}$ be different from zero for all $\mu \geqslant 0$ and let the sequence $[a, 0]_{k+1}[c, 0]_{k+1}^{-1}$, $[a, 0]_{k+2}[c, 0]_{k+2}^{-1}, \ldots$, be defined. Then the polynomials*

$$V_m(x) = (-1)^m[a, k+1]_m[c, m+k]_m^{-1}q^{\frac{1}{2}m(m-1)}F([\sigma, -m], [c, m+k]; [a, k+1]; xq)$$

are monic, orthogonal with respect to the sequence and have norms

$$p_{m,m+k} = [\sigma, 1]_m[a, 0]_{m+k+1}\{[c, 0]_{2m+k+1}[c, m+k]_m\}^{-1}\{[c, 0] - [a, 0]\} \ldots \\ \ldots \{[c, m-1] - [a, 0]\}\, q^{m(m+k)}.$$

Moreover we have the recurrence relation

$$V_{m+1}(x) = (x - a_m)\, V_m(x) - b_m V_{m-1}(x),$$

where

$$a_m = \frac{[\sigma, m]\{[c, m-1] - [a, 0]\}}{[c, 2m+k-1]_2} q^{m+k} + \frac{[a, m+k+1][c, m+k]}{[c, 2m+k]_2} q^m$$

and

$$b_m = [\sigma, m] \frac{[a, m+k][c, m+k-1]\{[c, m-1] - [a, 0]\}}{[c, 2m+k-2]_2[c, 2m+k-1]_2}.$$

If $[s, k]$ is given by (8.4), and we take $k = -1$, we obtain for $V_m(x)$:

$$V_m(x) = (-1)^m \frac{a(a+1)\ldots(a+m-1)}{(c+m-1)\ldots(c+2m-2)} F(-m, c+m-1; a; x),$$

which is easily seen to be a Jacobi polynomial (**9**), though in a different notation. For other values of k we get the same set of polynomials.

Hence, if $[s, k]$ is given by (8.5), we get Heine-analogues of the Jacobi polynomials.

There is a companion theorem to Theorem 13.1 corresponding to Theorem 6.4. This leads to the stepline polynomials

$$W_{2m}(x) = (-1)^m \frac{[a, k+1]_m}{[c, m+k]_m} q^{\frac{1}{2}m(m-1)}F([\sigma, -m], [c, m+k]; [a, k+1]; x^2q),$$

$$W_{2m+1}(x) = (-1)^m \frac{[c, m+k+1]_m}{[a, k+2]_m} q^{\frac{1}{2}m(m-1)}xF([\sigma, -m], [c, m+k+1]; \\ [a, k+2]; x^2q),$$

orthogonal with respect to $[a, 0]_{k+1}[c, 0]_{k+1}^{-1}$, 0, $[a, 0]_{k+2}[c, 0]_{k+2}^{-1}, 0, \ldots$, which in the case of ordinary hypergeometric series reduce to ultraspherical polynomials. Recurrence relations and norms can be deduced from Theorems 6.4 and 11.1. The well-known relation between Jacobi and ultraspherical polynomials actually originates in (6.2) and (6.3): if for $k = -1$ we denote $V_m(x)$ by $J_m(a, c; x)$ and $W_m(x)$ by $P_m(a, c; x)$, then $P_{2m}(a, c; x) = J_m(a, c; x^2)$, $P_{2m+1}(a, c; x) = xJ_m(a+1, c+1; x^2)$ (cf., for example, **9**, (4.1.5)). The same is true for the relation between Laguerre and Hermite polynomials (cf. **9**, (5.6.1)).

II. In a similar way we obtain from Theorem 12.1 as diagonal polynomials

$$V_m(x) = (-1)^m[a, k+1]_m q^{-m(m+k)} {}_1F_1([\sigma, -m]; [a, k+1]; xq^{m+k+1}),$$

which are orthogonal with respect to

$$[a, 0]_{k+1}\, q^{-\frac{1}{2}k(k+1)}, [a, 0]_{k+2} q^{-\frac{1}{2}(k+1)(k+2)}, [a, 0]_{k+3} q^{-\frac{1}{2}(k+2)(k+3)}, \dots,$$

if $[a, \mu] \neq 0$ for all $\mu \geqslant 0$ and if the moment sequence is defined. They are generalized Laguerre polynomials, and the corresponding stepline polynomials are generalized Hermite polynomials.

III. Applying Theorem 6.1 to the theorem resulting from Theorem 11.1 by substitution of ζ for c, $x[\zeta, 0]$ for x(cf. §12), we obtain the polynomials

$$V_m(x) = (-1)^m[a, k+1]_m q^{\frac{1}{2}m(m-1)} {}_1F_1^*([\sigma, -m]; [a, k+1]; xq),$$

which are orthogonal with respect to $[a, 0]_{k+1}$, $[a, 0]_{k+2}$, $[a, 0]_{k+3}, \dots,$ if $[a, \mu] \neq 0$ for $\mu \geqslant 0$, $a \neq \zeta$ and the moment-sequence is defined. They are also generalized Laguerre polynomials, and the corresponding stepline polynomials are likewise generalized Hermite polynomials.

IV. If in Theorem 11.1 we replace $[a, k]$ by $q^k y^{-1} + [\sigma, k]$, x by xy and put $y = 0$ we obtain

$$V_m(x) = (-1)^m[c, m+k]_m^{-1} q^{m(m+k)} {}_2F_0([\sigma, -m], [c, m+k]; xq^{-k}).$$

These are orthogonal with respect to

$$q^{\frac{1}{2}k(k+1)}[c, 0]_{k+1}^{-1}, q^{\frac{1}{2}(k+1)(k+2)}[c, 0]_{k+2}^{-1}, q^{\frac{1}{2}(k+2)(k+3)}[c, 0]_{k+3}^{-1}, \dots,$$

if the moment-sequence is defined. These polynomials are generalizations of the Bessel polynomials, introduced by Krall and Frink **(3)**.

V. If in Theorem 11.1 we substitute ζ for a, $x[\zeta, 0]^{-1}$ for x, we obtain

$$V_m(x) = (-1)^m[c, m+k]_m^{-1} q^{\frac{1}{2}m(m-1)} {}_2F_0^*([\sigma, -m], [c, m+k]; xq)$$

orthogonal with respect to $[c, 0]_{k+1}{}^{-1}$, $[c, 0]_{k+2}{}^{-1}$, $[c, 0]_{k+3}{}^{-1}, \dots,$ if $c \neq \zeta$ and the moment sequence is defined. These polynomials are also generalizations of the Bessel polynomials.

VI. If to the Padé-theorem obtained in III we apply the substitutions described in IV, or if to the Padé-theorem obtained in IV we apply the substitutions described in III we obtain in either case

$$V_m(x) = (-1)^m q^{m(m+k)} {}_1F_0^{**}([\sigma, -m]; xq^{-k}).$$

These polynomials are in fact the Stieltjes-Wigert polynomials, and are orthogonal with respect to $q^{\frac{1}{2}k(k+1)}$, $q^{\frac{1}{2}(k+1)(k+2)}$, $q^{\frac{1}{2}(k+2)(k+3)}, \dots,$ if $q \neq 1$.

VII. If to the Padé-theorem obtained in II we apply the substitutions described in V or if to the Padé-theorem obtained in V we apply the substitutions described in II, we obtain in each case

$$V_m(x) = (-1)^m q^{-m(m+k)} {}_1F_0^*([\sigma, -m]; xq^{m+k+1}).$$

These are also Stieltjes-Wigert polynomials, and in fact are the same as those in VI but with q^{-1} replacing q. Hence they are orthogonal with respect to $q^{-\frac{1}{2}k(k+1)}, q^{-\frac{1}{2}(k+1)(k+2)}, q^{-\frac{1}{2}(k+2)(k+3)}, \ldots$, if $q \neq 1$.

The expressions for the stepline polynomials in cases II to VII can be found from those for the diagonal polynomials. And also the recurrence relations and norms in cases II to VII and for the corresponding stepline polynomials can easily be deduced as in I.

PART III

14. Remarks concerning further generalization. The result of Part II being a generalization of the classical orthogonal polynomials, we note that the generalization of the hypergeometric series and consequently that of the classical orthogonal polynomials is strongly restricted by the rather unnatural-looking condition (8.8). This condition, it appears, has been introduced mainly to establish the recurrence relations (8.16) and (8.17), which are consequences of one another, and play an essential role in the proof of the key theorem 11.1 as they render the induction possible. It would therefore be much more natural to require, instead of (8.8), the existence of a relation of type (8.16). In view of (6.3), (11.6), and (11.8) the condition

$$(14.1)\quad F([a,k],[b,l];[c,m];x) - F([a,k+1],[b,l];[c,m+1];x) + \\ + \psi(a,b,c;k,l,m)x\, F([a,k+1],[b,l+1];[c,m+2];x) = 0,$$

where ψ is a suitably chosen function, looks very natural and general. In the following sections, however, it will be shown that the class of series satisfying (8.1), (8.2), and (14.1) is not essentially more general than that considered in Part II, in the sense that it does not give rise to a more general class of orthogonal polynomials.

From (14.1) it follows (cf. the proof of (8.16)) that

$$(14.2)\quad [a,k][c,m+\mu] - [a,k+\mu][c,m] \\ + \psi(a,b,c;k,l,m)[c,m][c,m+1][b,l]^{-1}[\sigma,\mu] = 0$$

for $\mu = 1, 2, 3, \ldots$. This can be written as

$$(14.3)\quad [a,k+\mu][c,m] - [a,k][c,m+\mu] = \chi(a,c;k,m)[\sigma,\mu], \mu = 1,2,3,\ldots$$

and hence the first problem is to find the functions $[s,k]$ and χ satisfying (14.3). We shall solve this problem in the next section. To avoid unessential difficulties we shall assume that R is algebraically closed. Since in the following a and c will be considered constant, we put $[a,k] = f(k)$, $[c,k] = g(k)$, $[\sigma,k] = h(k)$, $\chi(a,c;k,m) = \chi(k,m)$. Then the difference equation (14.3) becomes

$$(14.4)\quad f(k+\mu)\,g(m) - f(k)\,g(m+\mu) = \chi(k,m)\,h(\mu) \qquad \mu = 1,2,3,\ldots.$$

15. Solution of the difference equation. We shall first prove the following:

https://doi.org/10.4153/CJM-1958-060-2 Published online by Cambridge University Press

THEOREM 15.1. *The functions $f(k)$ and $g(k)$ satisfying* (14.4), *none of them being identically zero, satisfy the same trinomial linear recurrence relation with constant coefficients not all zero, that is,*

$$\begin{cases} f(k+2) + q f(k+1) + r f(k) = 0 \\ g(k+2) + q\, g(k+1) + r\, g(k) = 0 \end{cases} \qquad \text{for all } k. \tag{15.1}$$

Proof. First consider the case when f and g satisfy the relation

$$f(k+1)\ g(m) - f(k)\ g(m+1) = 0 \qquad \text{for all } k \text{ and } m \tag{15.2}$$

(this occurs, for example, if $\chi(k, m) \equiv 0$). There exist k_0 and m_0 such that $f(k_0) \neq 0$, $g(m_0) \neq 0$; hence, putting

$$f(k_0 + 1) f(k_0)^{-1} = g(m_0 + 1)\, g(m_0)^{-1} = p,$$

we have

$$f(k+1) - p f(k) = 0 \qquad \text{for all } k \tag{15.3}$$

$$g(m+1) - p\, g(m) = 0 \qquad \text{for all } m, \tag{15.4}$$

which are identical *binomial* linear recurrence relations.

Hence we may confine ourselves to the case when $\chi(k, m) \not\equiv 0$, and without loss of generality we may assume $\chi(0, 0) \neq 0$.

From (14.4) with $k = 0$, $m = 0$, we have

$$h(\mu) = \chi(0, 0)^{-1} \{f(\mu)\, g(0) - f(0)\, g(\mu)\}, \tag{15.5}$$

hence

$$f(k+\mu)g(m) - f(k)\ g(m+\mu) = \chi(k, m)\, \chi(0, 0)^{-1} \{f(\mu)g(0) - f(0)g(\mu)\}. \tag{15.6}$$

Substituting $\mu = 1$ resp. $\mu = 2$ we obtain

$$f(k+1)g(m) - f(k)g(m+1) = \chi(k, m)\, \chi(0, 0)^{-1} \{f(1)g(0) - f(0)g(1)\}, \tag{15.7}$$

$$f(k+2)g(m) - f(k)g(m+2) = \chi(k, m)\, \chi(0, 0)^{-1} \{f(2)g(0) - f(0)g(2)\}. \tag{15.8}$$

Elimination of $\chi(k, m)\, \chi(0, 0)^{-1}$ from (15.7) and (15.8) gives

$$\begin{aligned} &\{f(k+2)g(m) - f(k)g(m+2)\}\{f(1)g(0) - f(0)g(1)\} \\ &= \{f(k+1)g(m) - f(k)g(m+1)\}\{f(2)g(0) - f(0)g(2)\}. \end{aligned} \tag{15.9}$$

Substituting $m = 0$ in (15.9) and rearranging:

$$\begin{aligned} &f(k+2)g(0)\{f(1)g(0) - f(0)g(1)\} \\ &- f(k+1)g(0)\{f(2)g(0) - f(0)g(2)\} + f(k)g(0)\{g(1)f(2) - g(2)f(1)\} = 0. \end{aligned} \tag{15.10}$$

Substituting $k = 0$ in (15.9) and rearranging:

$$\begin{aligned} &g(m+2)f(0)\ \{f(1)g(0) - f(0)g(1)\} \\ &- g(m+1)f(0).\ \{f(2)g(0) - f(0)g(2)\} + g(m)f(0)\{g(1)f(2) - g(2)f(1)\} = 0. \end{aligned} \tag{15.11}$$

Without loss of generality we may assume that $f(1)g(0) - f(0)g(1) \neq 0$, since if $f(1)g(0) - f(0)g(1) = 0$, the formula (15.7) coincides with (15.2), a case which has already been considered.

We distinguish two possibilities. First suppose $f(k)g(m)\,\chi(k, m) \not\equiv 0$. Then we may assume without loss of generality that $f(0)g(0)\,\chi(0, 0) \neq 0$. And dividing (15.10) and (15.11) by $g(0)$ and $f(0)$ respectively it is clear that $f(k)$ and $g(k)$ satisfy the same trinomial linear recurrence relation with constant coefficients which are not all zero.

Now suppose $f(k)\,g(m)\,\chi(k, m) \equiv 0$. We still suppose $f(1)\,g(0) - f(0)\,g(1) \neq 0$. Now, if k_1 and m_1 are such that $f(k_1)$ and $g(m_1)$ are not zero, we have $\chi(k_1, m_1) = 0$ and hence

$$f(k_1 + \mu)\; g(m_1) - f(k_1)\, g(m_1 + \mu) = 0 \qquad \text{for } \mu = 1, 2, 3, \ldots .$$

Hence

$$g(m_1 + \mu) = g(m_1) f(k_1)^{-1} f(k_1 + \mu) \qquad \text{for } \mu = 1, 2, 3, \ldots .$$

From this it follows that $g(m)$ satisfies any linear recurrence relation with constant coefficients that is satisfied by $f(k)$. However, for $m = m_1$ it follows from (15.9) that $f(k)$ satisfies a trinomial linear recurrence relation with constant coefficients. This completes the proof of the theorem.

A check of this proof shows that (14.4) has only been used for $\mu = 1$ and 2.

With the assumption that R is algebraically closed, it is now easy to derive from Theorem 15.1 the functions f, g, h and χ satisfying (14.4). In fact, from the theory of linear difference equations, it follows (and it is easy to verify this directly), that any functions f and g satisfying (15.1) are given by

$$(15.12) \qquad f(k) = \omega_1 p_1^k + \omega_2 p_2^k, \quad g(m) = \omega_3 p_1^m + \omega_4 p_2^m ,$$

whenever the equation $x^2 + qx + r = 0$ has distinct roots p_1 and p_2. If this equation has two coincident roots p we have

$$(15.13) \qquad f(k) = (\omega_1 k + \omega_2)p^k, \quad g(m) = (\omega_3 m + \omega_4)p^m.$$

In both cases $\omega_1, \ldots, \omega_4$ are arbitrary constants.

Corresponding to (15.12) we find $\chi(k, m) = \omega_1\omega_4 p_1{}^k p_2{}^m - \omega_2\omega_3 p_2{}^k p_1{}^m$ and $h(\mu) = p_1{}^\mu - p_2{}^\mu$.

Corresponding to (15.13) we find

$$\chi(k, m) = \{\omega_1(\omega_3 m + \omega_4) - \omega_3(\omega_1 k + \omega_2)\}p^{k+m} \text{ and } h(\mu) = \mu p^\mu.$$

Recalling the role of the functions f, g, and h, we obtain the following:

THEOREM 15.2. *Every function* $[s, k]$ *such that the general hypergeometric series satisfies the functional equation* (14.1) *is given either by*

$$(15.14) \qquad [s, k] = \omega_1(s)p_1^k + \omega_2(s)p_2^k, \qquad \omega_1(\sigma) = -\omega_2(\sigma) = 1$$

or by

$$(15.15) \qquad [s, k] = \{k\,\omega_1(s) + \omega_2(s)\}p^k, \qquad \omega_1(\sigma) = 1, \quad \omega_2(\sigma) = 0,$$

where p, p_1 *and* p_2 *are not zero*, $p_1 \neq p_2$.

16. The nature of the more general series. Let us first consider the case in which $[s, k]$ is given by (15.14). Put $s' = -\omega_2(s)\,\omega_1(s)^{-1}$ if $\omega_1(s) \neq 0$, put $q = p_2 p_1^{-1}$, and put $[s', k]' = 1 - s'q^k$. Then, if $\omega_1(a)$, $\omega_1(b)$, $\omega_1(c)$ are all $\neq 0$, we find after some calculations that the series in (8.15) is equal to $F([a', k]', [b', l]'; [c', m]'; x\omega_1(a)\omega_1(b)\omega_1(c)^{-1}p_1^{k+l-m-1})$. Now consider the case that not all of $\omega_1(a)$, $\omega_1(b)$, $\omega_1(c)$ are different from zero, for example, $\omega_1(a) = 0$. Then the series in (8.15) is equal to ${}_1F_1([b', l]'; [c', m]'; x\omega_2(a)\,\omega_1(b)\omega_1(c)^{-1}p_2^k p_1^{l-m-1})$. Similar results are obtained in the other cases when one or more of $\omega_1(a)$, $\omega_1(b)$, $\omega_1(c)$ are zero. It follows that if $[s, k]$ is given by (15.14) the general hypergeometric series always coincides with a (possibly confluent) Heine series in which the variable x may be multiplied by a constant factor.

Let us now consider the case in which $[s, k]$ is given by (15.15). Put $s' = \omega_2(s)\omega_1(s)^{-1}$ if $\omega_1(s) \neq 0$ and put $[s', k]'' = s' + k$. If $\omega_1(a)$, $\omega_1(b)$ and $\omega_1(c)$ are different from zero, we find that the series in (8.15) is equal to

$$F([a', k]'', [b', l]''; [c', m]''; x\omega_1(a)\omega_1(b)\omega_1(c)^{-1}\, p^{k+l-m-1}).$$

Similarly, if one or more of the $\omega_1(a)$, $\omega_1(b)$, $\omega_1(c)$ are zero, we get confluent series. Investigation of all possible cases finally gives

THEOREM 16.1. *If the function $[s, k]$ is such that the series $F([a, k], [b, l]; [c, m]; x)$ satisfies the functional equation* (14.1), *then this series is always a (possibly confluent) ordinary hypergeometric or Heine series in which the variable x may be multiplied by a constant factor.*

This theorem implies the justification of our assertion in §14, that the series which satisfy (14.1) instead of (8.8) do not give a more general class of orthogonal polynomials than those obtained in Part II.

REFERENCES

1. E. Frank, *Orthogonality properties of C-fractions*, Bull. Amer. Math. Soc., *55* (1949), 384–390.
2. W. Hahn, *Ueber Orthogonalpolynome die q-Differenzen-gleichungen genügen*, Math. Nachr., *2* (1949), 4–34.
3. H. L. Krall, *Certain differential equations for Tchebycheff polynomials*, Duke Math. J., *4* (1938), 705–718.
4. H. L. Krall and O. Frink, *A new class of orthogonal polynomials: the Bessel polynomials*, Trans. Amer. Math. Soc., *65* (1949), 100–115.
5. H. Padé, *Récherches sur la convergence des développements en fractions continues d'une certaine catégorie des fonctions*, Ann. Sci. Ecole Norm. Sup., 3, *24* (1907), 341–400.
6. O. Perron, *Die Lehre von den Kettenbrüchen*, 2. Aufl. (Leipzig, 1929).
7. H. van Rossum, *A theory of orthogonal polynomials based on the Padé-table* (Diss. Utrecht, Assen 1953).
8. ———, *Systems of orthogonal and quasi-orthogonal polynomials connected with the Padé-table.* I, Kon. Ned. Akad. Wet., Proc. Section of Sciences, *58* Ser. A (1955), 517–525; II, *idem*, 526–534; III, *idem*, 675–682.
9. G. Szegö, *Orthogonal polynomials*, A.M.S. Coll. Publ. XXIII (New York, 1939).

University of New Brunswick