Pick's Theorem — What's the Big Deal?

John E. M$^{\text{c}}$Carthy

1 INTRODUCTION

In 1916, Georg Pick considered the following problem. Suppose that you are given points $\lambda_1, \ldots, \lambda_N$ in the unit disk of the complex plane (we shall denote the disk by $\mathbb{D}$); and suppose, too, that you are given complex numbers $w_1, \ldots, w_N$. Does there exist a holomorphic function $\phi : \mathbb{D} \to \mathbb{D}$ that interpolates, i.e., satisfies

$$\phi(\lambda_i) = w_i, \quad i = 1, \ldots, N? \tag{1}$$

Pick completely answered the question [8].[1]

Theorem 1 *A necessary and sufficient condition for the existence of a holomorphic function $\phi : \mathbb{D} \to \mathbb{D}$ satisfying the interpolation conditions (1) is that the self-adjoint N-by-N matrix*

$$\left[\frac{1 - \overline{w}_i \, w_j}{1 - \overline{\lambda}_i \, \lambda_j} \right] \tag{2}$$

be positive semi-definite. Moreover, ϕ is unique if and only if the Pick matrix (2) is singular, say of rank $M < N$; in this event, ϕ is a rational function that is an M-to-1 cover of $\mathbb{D}$.

It is obvious that one can always find some holomorphic function on $\mathbb{D}$ that interpolates (e.g., a polynomial); the restriction is whether a function can be found whose range is contained in $\mathbb{D}$. Let us rephrase this. Let $H^\infty(\mathbb{D})$

---

[1]Actually Pick considered the problem of interpolating w_i's in the right half-plane with a function of positive real part on $\mathbb{D}$. But because the right half-plane and unit disk are conformally equivalent, solving one problem solves the other.

denote the space of bounded holomorphic functions on $\mathbb{D}$, and define a norm on $H^\infty(\mathbb{D})$ by

$$\|\phi\|_{H^\infty(\mathbb{D})} \;=\; \sup_{z \in \mathbb{D}} |\phi(z)|.$$

Then Pick's problem is equivalent to the following:

(IP) *Among all bounded functions ϕ that satisfy the interpolation condition (1), what is the infimum of the $H^\infty(\mathbb{D})$-norms?*

Of course, finding the infimum is computationally harder than answering the question of whether the infimum is less than or equal to some fixed value, but mathematically the problems are equivalent. By a standard compactness argument (what a complex analyst would call a normal family argument), the infimum in (IP) will be attained. By rescaling, we can assume this infimum is 1, and then the Pick matrix (2) cannot be strictly positive definite (for if it were, it would remain so if all the w_i's were increased by some factor slightly bigger than one, which would mean by the first part of Pick's theorem that the infimum in (IP) was actually less than 1 to start with). The second part of Pick's theorem then says that the extremal function is unique, and is a constant times a *Blaschke product,* i.e., it has the form

$$\phi(z) \;=\; c \prod_{i=1}^{M} \frac{z - \zeta_i}{1 - \bar{\zeta_i} z}$$

for some constant c and some points $\zeta_1, \ldots, \zeta_M$ in $\mathbb{D}$. For any fixed ζ in $\mathbb{D}$, the function

$$\psi(z) \;=\; \frac{z - \zeta}{1 - \bar{\zeta}}$$

is an automorphism of $\mathbb{D}$, and thus maps the boundary $\partial\mathbb{D}$ to $\partial\mathbb{D}$. It is called a *Blaschke factor* (with zero at ζ). A product of Blaschke factors therefore has modulus one everywhere on $\partial\mathbb{D}$, and indeed all rational holomorphic functions on $\mathbb{D}$ that have modulus one everywhere on $\partial\mathbb{D}$ must be a finite product of Blaschke factors times a constant of modulus one. Thus extremal functions, which are constrained to have modulus less than or equal to $|c|$ everywhere, must have modulus identically $|c|$ on $\partial\mathbb{D}$.

Because of the First World War, Rolf Nevanlinna in Finland was unaware of Pick's result, even though it was published in *Mathematische Annalen.* He considered the same problem in [6], and analyzed it using an idea of Issai Schur [10], [11], resulting in a different characterization. In 1929 Nevanlinna

gave a parametrization of all solutions in the nonunique case [7]. Since then, this type of problem has been called *Pick interpolation* or *Nevanlinna-Pick interpolation*.

A search of MathSciNet reveals that in the last decade alone, over 300 articles have been written on Pick interpolation. Why so much interest in a problem that was completely solved so long ago that almost everyone would deem it classical?[2]

The purpose of this article is to explain why the problem comes up in engineering, and what the modern[3] approach to it is.

2   CONTROL THEORY

In control theory, one studies some physical system, called a plant, that has an input and an output. The plant could be the steering mechanism of a boat, with the input the signal from either a human at the wheel or an autopilot, and the output the directions to the engines and rudder. Or it could be a thermostat in a house, with input the temperature at various points in the house, and output instructions to the heating and cooling systems. In practice, one does not pay too much attention to the internal workings of the plant. One just assumes that, either by knowledge of the plant's construction or by experimental measurement, one knows how the plant responds to the input signals.

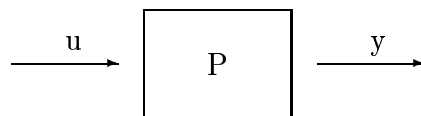A basic system is therefore just a plant, which we treat as a black box:

Figure 1: The original system

---

[2]A result is called *classical* if it was proved before the speaker went to graduate school. Not to be confused with *classic*, a result that the speaker has improved upon.

[3]The only approach the speaker really understands.

We shall use discrete time for convenience, so the input is a sequence $\{u_n\}_{n=0}^{\infty}$ in the vector space $\mathbb{C}^r$, with energy $\sum_{n=0}^{\infty}\|u_n\|^2$, where for any vector $x = (x_1, x_2, \ldots, x_r)$ in $\mathbb{C}^r$ we define $\|x\|^2 = |x_1|^2 + |x_2|^2 + \ldots + |x_r|^2$. (In a continuous time model, the input would be a $\mathbb{C}^r$-valued function $u(t)$; we assume we measure this only at discrete times of separation h, and then u_n is just the value of $u(nh)$.) The output is a sequence $\{y_n\}_{n=0}^{\infty}$ in $\mathbb{C}^s$. For the time being, let us take $r = s = 1$.

The Plant. The plant P is assumed to have the following four properties:

(C) Causality $(u_n = 0 \,\forall\, n \leq k \Rightarrow y_n = 0,\ \forall\, n \leq k)$
(TI) Time Invariance   (Shift input $\rightarrow$ shift output)
(S) Stability   (energy doesn't increase indefinitely)
(L) Linearity   (!!)

The first three assumptions are all physically reasonable. (C) says that when no signal is coming in, the output is zero. (TI) says that, if the input $(u_0, u_1, u_2, \ldots)$ produces output $(y_0, y_1, y_2, \ldots)$, then the input $(0, u_0, u_1, \ldots)$ will produce output $(0, y_0, y_1, \ldots)$. (S) says that there is some constant M (often not larger than 1) such that for any input the energy of the output is no greater than M times the energy of the input.

Linearity (L) is the big deal. Locally, all (differentiable) systems are linear—this is the basic insight of differential calculus. But the region of linearity is much greater than one has a right to expect.

Fundamental Principle of Engineering. *Take any black box Plant, as in Figure 1. Then its behaviour is linear over all normal[4] inputs. Moreover, if the input signal is $\sin(\omega t)$, the output is $A\sin(\omega t + c)$.*

Obviously, we are over-generalizing. Transistors and many other objects are essentially nonlinear. Nevertheless, for most systems, it is an experimental fact that the principle holds. This can partly be explained mathematically by the realization that a collection of masses connected by damped springs, or an RCI electrical circuit, can be modelled as a system of second order linear ODE's. One does not need to know the details of the configuration of such a system to know that its behavior will be linear, and that it will be frequency preserving. So by measuring its response to a range of sig-

---

[4] A normal input is roughly one that is not so strong that it damages the plant, nor so weak that it is below the sensitivity of the plant.

nals of varying frequencies, one can predict its response to an input that is synthesized from these frequencies.

The z-transform. The z-transform sends the sequence $\{u_n\}$ to the analytic function $\tilde{u}$ whose Taylor coefficients at 0 are the terms u_n:

$$Z \; : \; \{u_n\}_{n=0}^{\infty} \; \mapsto \; \tilde{u}(z) \; := \; \sum_{n=0}^{\infty} u_n z^n. \tag{3}$$

If $u = \{u_n\}$ has finite energy (i.e., $\sum |u_n|^2 < \infty$), then $\tilde{u}(z) = \sum_{n=0}^{\infty} u_n z^n$ is analytic in the unit disk, and its energy is given by

$$\sup_{0 \le r < 1} \frac{1}{2\pi} \int_0^{2\pi} |\tilde{u}(re^{i\theta})|^2 d\theta \; = \; \lim_{r \nearrow 1} \sum_{n=0}^{\infty} r^{2n} |u_n|^2. \tag{4}$$

The space of analytic functions $\tilde{u}$ in $\mathbb{D}$ for which (4) is finite is called the *Hardy space*, and it forms a Hilbert space with (4) as the square of the norm. The Hardy space is denoted by H^2.

After identifying sequences with the functions that are their z-transforms, the correspondence $P : \tilde{u} \mapsto \tilde{y}$ of input to output defines a linear operator from H^2 to H^2. Assumption (S) means that P is bounded (which, for linear operators, is the same as being continuous). Suppose $P : 1 \mapsto \phi$. Then by (C), (TI), and (L), for any polynomial p we have

$$P : p \mapsto \phi p.$$

By continuity, P is then multiplication by ϕ on all of H^2.

What can we say about $\|P\|$, the operator norm of P, in terms of ϕ? As

$$\|P\|^2 \; = \; \sup_{\|f\|_{H^2} \le 1} \sup_{0 < r < 1} \frac{1}{2\pi} \int_0^{2\pi} |\phi(re^{i\theta}) f(re^{i\theta})|^2 d\theta, \tag{5}$$

we find immediately that

$$\|P\| \; \le \; \sup_{z \in \mathbb{D}} |\phi(z)|. \tag{6}$$

We actually have equality in (6). Indeed, suppose that the left-hand side is 1 (we can achieve this by scaling). Then, taking $f = 1$ in (5), we have

$$
\begin{aligned}
1 \;\; &\ge \;\; \|P^n 1\|^2 \\
&= \;\; \sup_{0 < r < 1} \frac{1}{2\pi} \int_0^{2\pi} |\phi(re^{i\theta})|^{2n} d\theta.
\end{aligned}
\tag{7}
$$

If the supremum of $|\phi(z)|$ were greater than 1, then there would be a set of length $\delta > 0$ where $|\phi| > 1 + \varepsilon$. Then (7) would be bounded below by $\frac{1}{2\pi}(1+\varepsilon)^{2n}\delta$, which tends to infinity with n.

Thus we have proved:

Theorem 2 *Under assumptions* (C) – (L), *the plant P is given by*

$$P : \tilde{u} \mapsto \phi\tilde{u} \tag{8}$$

for some ϕ in $H^\infty(\mathbb{D})$. Moreover

$$\|P\| = \sup_{z \in \mathbb{D}} |\phi(z)|.$$

From here on, we shall adopt the standard notation of using the same letter P for both the plant and the $H^\infty(\mathbb{D})$ function ϕ from (8).

The Compensator. Of course, real plants are subject to some unknown noise e, whose effects are modulated by some other plant W (which we assume is known), as depicted in Figure 2.
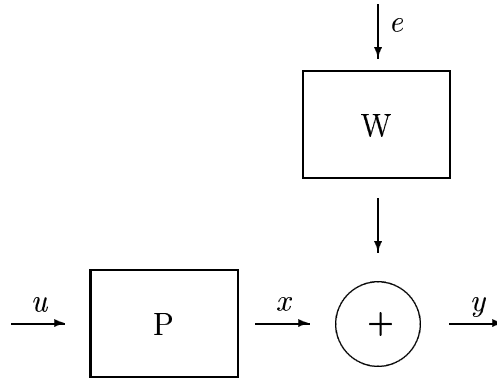

Figure 2:

So we introduce a feed-back loop and a compensator C, which we can design. We end up with the system shown in Figure 2, with inputs u (the signal) and e (the noise), and output y.
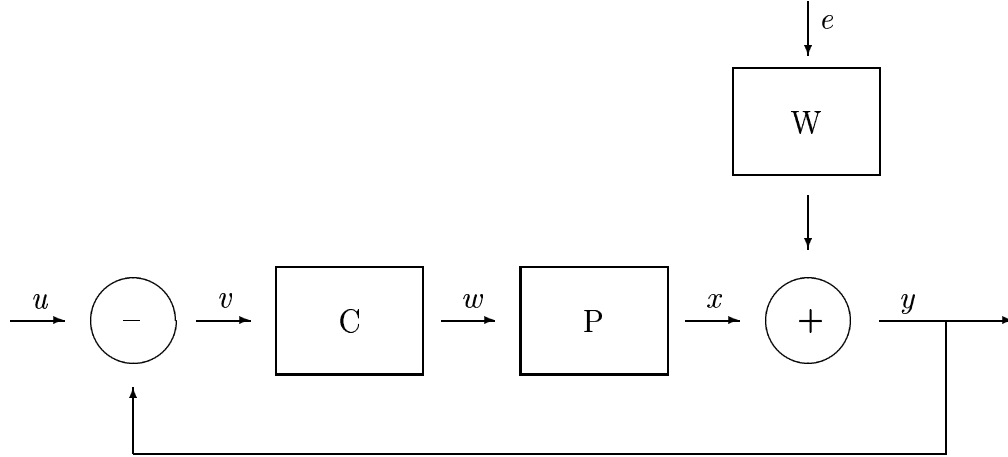
Figure 3:

Solving, we get

$$y = PC(u - y) + W\,e,$$

so

$$y = (I + PC)^{-1}\,PCu + (I + PC)^{-1}W\,e. \tag{9}$$

Internal stability. We assume that both P and W are stable (i.e., after taking the z-transform, they correspond to multiplication by $H^\infty(\mathbb{D})$ functions). We do not insist that C be stable, but we do require *internal stability*, meaning that the internal signals v, w, and x must have finite energy when the inputs u and e do. Because

$$
\begin{aligned}
v &= y - u \\
&= [(I + PC)^{-1}\,PC - I]u + (I + PC)^{-1}W\,e \\
&= -(I + PC)^{-1}\,PCu + (I + PC)^{-1}W\,e, \\
w &= Cv, \\
x &= Pw,
\end{aligned}
$$

internal stability requires that $(I + PC)^{-1}$ and $C(I + PC)^{-1}$ be stable (since P is already assumed stable and the product of stable plants is stable). The

7

identity
$$(I + PC)^{-1} = I - P[C(I + PC)^{-1}]$$
means that internal stability is equivalent to requiring that
$$F := C(I + PC)^{-1} \quad \text{be stable.} \tag{10}$$

Model-matching problem. We would like to design the compensator C to minimize the effect of the noise on the output. In view of (9), this means that we want to minimize
$$||(I + PC)^{-1}W|| = ||(I - PF)W||.$$
The constraint, from (10), is that F lie in $H^{\infty}(\mathbb{D})$.

Let us make a very mild simplifying assumption: we assume that both P and W are rational functions. Then they can be factored as
$$P = P_i P_o,$$
$$W = W_i W_o.$$

Here, P_i and W_i are Blaschke products, with zeroes precisely at the zeroes in the disk of P and W, respectively, and P_o and W_o are rational functions that have no zeroes in the open unit disk. (The Blaschke product is called the *inner factor* of the rational function, the other one is called the *outer factor*. Hence the subscripts i and o.) We want to find
$$\begin{aligned}
(*) &= \inf_{F \in H^{\infty}(\mathbb{D})} ||W - PFW|| \\
&= \inf ||W_o W_i - P_i P_o F W_o W_i|| \\
&= \inf ||W_o - P_i(P_o F W_o)||.
\end{aligned}$$

(The last equality comes from the maximum principle. The supremum of the modulus of an analytic function on $\mathbb{D}$ is the limit of the suprema over circles of radius r as r increases to 1. On these circles, the modulus of a Blaschke product tends to 1.)

If neither P_o nor W_o has zeroes on the boundary of $\mathbb{D}$, then as F ranges over $H^{\infty}(\mathbb{D})$, so does $G := P_o F W_o$. If one of the outer factors does have a zero on the boundary, then the set of attainable G's is still dense in an

appropriate topology (the weak-star topology—see [1] for a discussion). So our problem reduces to finding

$$(*) \;=\; \inf_{G \in H^\infty(\mathbb{D})} \|W_o - P_i G\|.$$

Let the zeroes of P_i be $\lambda_1, \ldots, \lambda_N$. Then

$$\{W_o - P_i G \,:\, G \,\in\, H^\infty(\mathbb{D})\}$$
$$=\; \{H \,\in\, H^\infty(\mathbb{D}) \,:\, H(\lambda_m) = W_o(\lambda_m),\ 1 \le m \le N\}.$$

Our problem thus translates into finding the function of smallest norm that interpolates N given values—the Pick problem!

For simplicity, we worked in the scalar case. In practice, engineers care about the matrix case, where the input u and output y are vectors—for example, y could be the position and velocity of a ship, and u could be the steering commands. A similar analysis of more complicated systems yields the following problem, called the *model-matching problem*, which is fundamental in control theory:

Model-matching Problem: *Given matrix-valued $H^\infty(\mathbb{D})$ functions Q, P, and W, find the analytic matrix-valued function F that minimizes $\|Q - PFW\|_\infty$.*

For a more detailed discussion of these problems, we refer the reader to [3] and the references therein.

3 THE OPERATOR THEORY APPROACH.

Whilst Pick's theorem can be proved using classical function theory (see, for example, [5]), it can also be approached via operator theory on the Hardy space H^2. Although we shall again stick to the scalar case for simplicity, one advantage of the operator theory approach is that it generalizes naturally to the matrix-valued case.

Recall that H^2 is the Hilbert space of all functions $f(z) = \sum_{n=0}^\infty \hat{f}(n) z^n$ that are analytic in $\mathbb{D}$ and have square-summable Taylor coefficients at the origin. The inner product is given by

$$\langle f, g \rangle \;=\; \sum_{n=0}^\infty \hat{f}(n) \overline{\hat{g}(n)}$$

9

$$= \lim_{r \nearrow 1} \frac{1}{2\pi} \int_0^{2\pi} f(re^{i\theta})\overline{g(re^{i\theta})}d\theta.$$

We proved in Section 2 that, if ϕ is in $H^\infty(\mathbb{D})$, then the operator of multiplication by ϕ, which we shall denote M_ϕ, is a bounded operator on H^2 and moreover

$$\|M_\phi\|_{B(H^2)} = \|\phi\|_{H^\infty(\mathbb{D})}.$$

So Pick's problem asks whether there exists a contractive multiplier satisfying certain interpolation conditions.

The Szegő kernel. For each λ in $\mathbb{D}$, define the function k_λ in H^2 by

$$k_\lambda(z) = \frac{1}{1 - \bar{\lambda}z} = \sum_{n=0}^\infty \bar{\lambda}^n z^n.$$

This is called the *Szegő kernel*. It is the reproducing kernel for H^2, in the following sense:

$$\langle f, k_\lambda \rangle = \sum_{n=0}^\infty \hat{f}(n)\lambda^n = f(\lambda),$$

for every f in H^2. Moreover, for every multiplier ϕ, the function k_λ is an eigenvector of the adjoint M_ϕ^*:

Lemma 1 *Let ϕ be in $H^\infty(\mathbb{D})$, and let λ be a point in $\mathbb{D}$. Then*

$$M_\phi^* k_\lambda = \overline{\phi(\lambda)}k_\lambda. \tag{11}$$

Proof: Let f be an arbitrary function in H^2. Then

$$\begin{aligned}
\langle f, M_\phi^* k_\lambda \rangle &= \langle \phi f, k_\lambda \rangle \\
&= \phi(\lambda)f(\lambda) \\
&= \langle f, [\overline{\phi(\lambda)}k_\lambda] \rangle.
\end{aligned}$$

So the two sides of (11) have the same inner product with every element f of H^2, and therefore they must be equal. □

The necessity of Pick's condition. Suppose that ϕ is a function in the closed unit ball of $H^\infty(\mathbb{D})$ that interpolates λ_i to w_i as in (1). Then

$\|M_\phi^*\| \leq 1$, so for any constants $c_1, \ldots, c_N$ we must have

$$\Big\langle M_\phi^*(\sum_{i=1}^{N} c_i k_{\lambda_i}), M_\phi^*(\sum_{j=1}^{N} c_j k_{\lambda_j}) \Big\rangle \leq \Big\langle \sum_{i=1}^{N} c_i k_{\lambda_i}, \sum_{i=1}^{N} c_j k_{\lambda_j} \Big\rangle. \qquad (12)$$

By (11),

$$M_\phi^* k_{\lambda_i} = \bar{w}_i k_{\lambda_i}.$$

Accordingly (12) becomes

$$\sum_{i,j=1}^{N} c_i \overline{c_j}(1 - \bar{w}_i w_j)\langle k_{\lambda_i}, k_{\lambda_j} \rangle \geq 0. \qquad (13)$$

Because

$$\langle k_{\lambda_i}, k_{\lambda_j} \rangle = \frac{1}{1 - \bar{\lambda}_i \lambda_j},$$

the positivity of (13) for all choices of scalars c_i is precisely the statement that the Pick matrix (2) be positive. Thus we have proved the necessity of Pick's condition in Theorem 1.

Sarason's idea. By running the preceding argument backwards, the assertion that the Pick matrix (2) is positive is seen to be equivalent to the statement:

(A.) *The linear operator R, defined on the N-dimensional subspace $\mathcal{M}$ of H^2 spanned by the kernel functions k_{λ_i}, $1 \leq i \leq N$, is a contraction.*

The converse of Pick's theorem is the assertion that, if (A.) holds, then

(B.) *R extends to an operator Y on all of H^2 that is contractive and is the adjoint of a multiplication operator.*

There is a distinguished operator on H^2 called the *unilateral shift* and denoted by S. It is defined by

$$[Sf](z) = zf(z),$$

so it shifts the Taylor coefficients of f to the right. (Notice that the definition of (TI) is exactly that the plant commutes with the shift.) Its adjoint S^* is called the *backward shift*.

Donald Sarason observed in [9] that condition (A.) can be rephrased as saying that R is a contraction on $\mathcal{M}$ that commutes with the restriction of S^*

to $\mathcal{M}$, while condition (B.) asserts that R can be extended to a contraction Y on H^2 that commutes with S^*. Sarason proved the following theorem, from which the sufficiency of Pick's condition follows immediately.

Theorem 3 *Let $\mathcal{M}$ be an invariant subspace for S^*, and let R be an operator on $\mathcal{M}$ that commutes with $S^*|_{\mathcal{M}}$. Then there exists an operator Y on H^2 that commutes with S^*, has the same norm as R, and satisfies $Y|_{\mathcal{M}} = R$.*

Sarason's theorem was generalized by Bela Sz.-Nagy and Ciprian Foiaş to their celebrated commutant lifting theorem [12], [13]. This theorem says roughly that Sarason's Theorem 3 remains true if S^* is replaced by the direct sum of arbitrarily many copies of S^*. From the commutant lifting theorem one can derive not only matrix-valued versions of Pick's theorem, but also solutions to various other model-matching problems. See [2] for a discussion.

4 CONCLUSION

Pick's problem is only a special case of the type of problems that arise in H^∞ control theory. However, it is both deep enough that solutions of the Pick problem can often be generalized to solutions of other problems, and transparent enough that one can find solutions to the Pick problem in different ways.

Many variants on Pick's problem are only partially solved. If one leaves the unit disk and replaces the domain of the functions by some other set, like an annulus or a polydisk, things get more complicated. See [1] for a discussion of some known results of this type. If one changes the range, and wants functions whose values all lie in an annulus, or in the intersection of two disks, then the problem gets even harder. We refer to [4] for a discussion of some of these problems.

Eighty-six years after its solution, Pick's problem is still inspiring new pure mathematics, and new applied mathematics!

References

[1] J. Agler and J.E. McCarthy, *Pick Interpolation and Hilbert Function Spaces*, American Mathematical Society, Providence, 2002.

[2] C. Foiaş and A.E. Frazho, *The Commutant Lifting Approach to Interpolation Problems*, Birkhäuser, Basel, 1990.

[3] J.W. Helton, *Operator theory, Analytic Functions Matrices and Electrical Engineering*, number 68 in CBMS Lecture Notes, American Mathematical Society, Providence, 1987.

[4] J.W. Helton and M.R. James, *Extending H^∞ Control to Nonlinear Systems. Control of Nonlinear Systems to Achieve Performance Objectives*, Advances in Design and Control, 1. Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[5] D. Marshall, An elementary proof of the Pick-Nevanlinna interpolation theorem, *Michigan Math. J.* 21 (1974) 219–233.

[6] R. Nevanlinna, Über beschränkte Funktionen, die in gegebenen Punkten vorgeschrieben Werte annehmen, *Ann. Acad. Sci. Fenn. Ser. A* 13(1), (1919) 1–72.

[7] R. Nevanlinna, Über beschränkte Funktionen, *Ann. Acad. Sci. Fenn. Ser. A*, 32(7), (1929)

[8] G. Pick, Über die Beschränkungen analytischer Funktionen, welche durch vorgegebene Funktionswerte bewirkt werden, *Math. Ann.*, 77 (1916) 7–23.

[9] D. Sarason, Generalized interpolation in H^∞, *Trans. Amer. Math. Soc.*, 127 (1967) 179–203.

[10] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind I, *J. Reine Angew. Math.*, 147 (1917) 205–232.

[11] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind II, *J. Reine Angew. Math.*, 148 (1918) 122–145.

[12] B. Sz.-Nagy and C. Foiaş, Commutants de certains opérateurs, *Acta Sci. Math. (Szeged)*, 29 (1968) 1–17.

[13] B. Sz.-Nagy and C. Foiaş, Dilatations des commutants d'opérateurs, *C. R. Acad. Sci. Paris Sér. A-B*, 266 (1968) A493–A495.

John Edward McCarthy is a professor of mathematics at Washington University. He was an undergraduate at Trinity College, Dublin, and got a Ph.D. from U.C. Berkeley. His papers are available at *www.math.wustl.edu/ ˜mccarthy/papers.html.* His hobbies are doing mathematics, drinking coffee and saving the world.
Washington University, Saint Louis, MO 63130
mccarthy@math.wustl.edu