Euler, Stirling, and Wallis: A Case Study in the Notion of Equivalence Between Theorems

Fausto di Biase (Università "G. D'Annunzio" di Chieti-Pescara, Pescara, Italy)

The notion that two theorems may be equivalent to each other is sometimes met with hesitation. In this article we tell a story that shows that there is something interesting and useful in this notion. We look at the following three results: Stirling's formula, Wallis' product formula, and the evaluation of the probability integral. The task of giving simple *proofs of these results is the object of unabated attention. In order to enhance our understanding of these results, we show in a precise way that these results are indeed* equivalent *to each other.*

1 Introduction

A controversial notion

The notion that two theorems may be equivalent is sometimes met with hesitation, since *a posteriori* i.e., *after* they have been proved, there is apparently not much more to say.

Example 1. Consider the following two statements:
- "The Fourier series of an L^2 function converges almost everywhere";
- "The maximal operator associated to the Fourier series is of weak type (2,2)".

Before the truth value of any of the two statements had been determined, in the mid-fifties it was understood that they are equivalent to each other – thanks to the work of A. Calderón and A. Zygmund, based on previous work of G. H. Hardy, and J. E. Littlewood.[1] This equivalence – whose roots lie in a work by Kolmogorov [8] – turned out to be very useful, for two reasons. Firstly, it pointed to a specific research agenda. Secondly, it was a hint that perhaps hidden in the background there was some more general result – which indeed was discovered a few years later by E. M. Stein in a seminal, important work [16]. This example shows how the equivalence between two *statements* can be useful if it can be established *before* the two statements have been proven.

Example 2. Consider the following two statements:
- Every continuous map from the closed disc into itself has at least a fixed point;
- The Hex game cannot end in a draw.

These two statements are equivalent [4]. As a matter of fact, each of these two statements is a theorem. In this example, the equivalence is considered of interest – even if it was establishe *after* the two results had been proved – since it points to a surprising connection between two apparently remote topics.

We would like to tell a story that will perhaps dispel the diffidence toward this notion.

A general viewpoint

The general viewpoint that we would like to suggest in this paper is that a proof that two statements are equivalent will enhance our understanding of the subject, provided their equivalence assumes the form of a *theorem* which represents an *a priori, direct link* between the two statements, where
- the term *a priori* means that the connection is not contingent on the truth value of the two statements – a notion vaguely inspired by modal logic [9];
- the theorem is a *direct link* in the sense that it shows *at once*, in a visible and immediate way, and *without any further substancial work*, that the two statements must *necessarily* have the same truth value.

We should emphasize the fact that an a priori, direct link between two statements is *a theorem* – a positive result which adds something to our knowledge, not a purely conjectural result. In particular, even if the two equivalent statements at hand should one day turn out to be false, the a priori, direct link which we have determined will have increased our knowledge.

Now, we would like to give an example of an *a priori, direct link* between two well-known results. Firstly, recall that if F and G are sequences of positive real numbers, then the symbol $F_n \sim G_n$ means that $\lim_{n \to +\infty} \frac{F_n}{G_n} = 1$.

Example 3. Consider the following two statements:

$$n! \sim e^{-n} n^n \sqrt{2\pi n} \qquad \text{(S)}$$

$$\frac{(2n)!!}{(2n-1)!!} \sim \sqrt{n\pi} \qquad \text{(W)}$$

The following result is an a priori, direct link between (S) and (W).

Theorem 4 (Moritz [13])**.**

$$n! \sim \frac{(2n)!!}{(2n-1)!!} e^{-n} n^n \sqrt{2} \qquad \text{(W↔S)}$$

The tag (↔) chosen for this formula alludes to the fact that it shows *at once*, in a visible and immediate way, and *without any further substancial work*, that the two statements must *necessarily* have the same truth value. Moreover, the statement in Theorem 4 is not contingent on the truth value of the two statements (S) and (W). Hence, Theorem 4 is an a priori, direct link between (S) and (W). Theorem 4 is enlightening because it shows that (S) and (W) are different sides of the same object, since it points to a *necessary* connection between their truth values: It contains the heart of the connection between the two statements.

---

1 [19, Theorem 1.22, v.2]

The result in (S) is due to James Stirling (1692–1770), who obtained it while working on problems of probability theory by completing a previous result by Abraham De Moivre (1667–1754). It is very important in probability theory and it has deep connections to many other topics.

The result in (W) was proved by John Wallis (1616–1703) in Proposition 191 of his book *The Arithmetic of Infinitesimals*, published in 1656 [17]. Wallis was interested in the application of the new ideas of infinitesimal calculus to the computation of areas, in particular, the area of a circle, and was led to this result by a clever use of the idea of interpolation – an idea that was later used by Euler with great virtuosity. Wallis' book had a great influence on Newton, although it had been met with sharp criticism by Fermat, Huygens and Hobbes.

The probability integral
This name refers to the following result (due to Euler [2], [3]).

$$\int_{\mathbb{R}} e^{-x^2}\, dx = \sqrt{\pi} \qquad \text{(E)}$$

It was rediscovered by Laplace [10], [11] and played a fundamental role in the work of Gauss [5]. Indeed, it plays an important role in the method that Gauss used to predict the position of Ceres on the basis of very few observations [6].

A tale of three stories
The task of giving *simple* proofs of (S), (W), and (E) is the object of unabated attention. There are currently many *simple* proofs of these results, of which [7], [12], [13], [14], [18] are just a small sample.

Below we illustrate the general viewpoint of Sec. 1 by looking for a priori, direct links between these results.[2] In particular, our main intent is *not* to obtain yet another simple proofs of these results.

We have already seen that Theorem 4 looks at the first two results – (S) and (W) – from the general viewpoint of Sec. 1, and establishes an a priori, direct link between them. In this paper we try to do for the pair of statements (E) and (W) what Moritz did for (S) and (W).

2 A priori, direct links between (E) and (W)

Lemma 5. *If l_n is any sequence of positive numbers such that $\lim_{n \to +\infty} l_n = +\infty$ then*

$$\int_{\mathbb{R}} e^{-x^2}\, dx = \lim_{n \to +\infty} \int_{-l_n}^{l_n} \left(1 + x^2/n\right)^{-n} dx \qquad (1)$$

Proof. Observe that $e^{-x^2} \le \left(1 + x^2/n\right)^{-n} \le \left(1 + x^2\right)^{-1}$. Indeed, $e^z \le (1 - z)^{-1}$ for $z \le 0$, then let $z = -x^2/n$. The sequence $\left(1 + x^2/n\right)^{-n}$ converges uniformly to e^{-x^2} in each bounded interval, as can be seen by the Taylor series of ln

---

2 We will try to use in our proofs as little machinery as we can and, moreover, we will try to offer proofs conforming to the standards of *clarity* described in [1]. An example *in the negative* is given by proofs by contradiction, which are opaque, as observed in [15, p.3], since they shed little insight into the link between hypothesis and conclusion.

and e^x. Now, given $\epsilon > 0$, choose $c > 0$ such that

$$I := \int_{\mathbb{R} \setminus [-c,c]} e^{-x^2}\, dx < \epsilon$$

and

$$II := \int_{\mathbb{R} \setminus [-c,c]} \left(1 + x^2\right)^{-1} dx < \epsilon$$

as we may, due to the absolute integrability of the functions. If n is large enough, then $l_n > c$ and

$$III := \int_{[-c,c]} \left| \left(1 + x^2/n\right)^{-n} - e^{-x^2} \right| dx < \epsilon$$

due to uniform convergence, thus

$$\left| \int_{[-l_n,l_n]} \left[\left(1 + x^2/n\right)^{-n} - e^{-x^2} \right] dx \right| \le I + II + III < 3\epsilon \quad \square$$

Corollary 6. *There is an a priori, direct link between* (E) *and the following statement*

$$\lim_{n \to +\infty} \sqrt{n} \int_0^{\pi} (\sin \theta)^{2n}\, d\theta = \sqrt{\pi} \qquad \text{(E.1)}$$

Proof. It suffices to show that

$$\int_{\mathbb{R}} e^{-x^2}\, dx = \lim_{n \to +\infty} \sqrt{n} \int_0^{\pi} (\sin \theta)^{2n}\, d\theta \qquad (\text{E} \leftrightarrow \text{E.1})$$

Let $l_n = \sqrt{n}$ in Lemma 5 and let $\frac{x}{\sqrt{n}} = z$ and $z = \cot \theta$ in this order. Then

$$\int_{-\sqrt{n}}^{\sqrt{n}} \left(1 + x^2/n\right)^{-n} dx = \sqrt{n} \int_{-1}^{1} \left(1 + z^2\right)^{-n} dz$$

$$= \sqrt{n} \int_{\pi/4}^{3\pi/4} (\sin \theta)^{2n-2}\, d\theta$$

Now observe that

$$\sqrt{n} \int_{\pi/4}^{3\pi/4} (\sin \theta)^{2n-2}\, d\theta = \sqrt{n} \int_0^{\pi} (\sin \theta)^{2n-2}\, d\theta + o(1)$$

since the difference is bounded by $\pi \sqrt{n} 2^{-n}$. Finally, integration by parts yields the familiar reduction formula

$$n \cdot S_n = (n - 1) \cdot S_{n-2} \quad (n \ge 2) \qquad (2)$$

for the sequence $S = \{S_n\}_n$ defined by

$$S_n = \int_0^{\pi} (\sin \theta)^n\, d\theta \qquad (3)$$

which implies that

$$S_{2n} \sim S_{2n-2} \qquad (4)$$

We have thus proved that $\int_{\mathbb{R}} e^{-x^2}\, dx = \lim_{n \to +\infty} \sqrt{n} S_{2n}$, hence the conclusion. $\square$

Corollary 7. *The statement in* (E) *is equivalent to*

$$S_{2n} \sim \sqrt{\frac{\pi}{n}} \qquad \text{(E.2)}$$

Observe that the second order homogeneous difference equation (2), together with the initial values S_0 and S_1, uniquely determines the sequence S. A more general point of view yields a better view. Let $\mathbb{S} = (0, +\infty)^{\mathbb{N}}$ be the set of all sequences of strictly positive numbers. If $F = \{F_n\}_n \in \mathbb{S}$ and $p, q \in (0, +\infty)$, then we define $F^*(p,q) = \{F_n^*(p,q)\}_n \in \mathbb{S}$ as

the unique solution of the second order homogeneous difference equation

$$F(n)F_n^*(p,q) = F(n-1) \cdot F_{n-2}^*(p,q) \quad (n \geq 2) \qquad (5)$$

with initial conditions $F_0^*(p,q) = p$, $F_1^*(p,q) = q$. The assignment $F \mapsto F^*(p,q)$ is a map $\mathbb{S} \to \mathbb{S}$: This map is *surjective* but *not* injective. We call $F^*(p,q)$ the sequence *associated to* F, with initial data (p,q). If a statement does not depend on the special choice of the initial data, we denote $F^*(p,q)$ by F^*.

Example 8. Let $I = \{I_n\}_n \in \mathbb{S}$ be the *identity sequence*, defined by $I_n = n$. Since $S_0 = \pi$ and $S_1 = 2$, then (2) says that the sequence S defined in (3) is the sequence $I^*(\pi, 2)$ associated to I with initial data π and 2.

Remark 9. Observe that the sequence I, which enters in the reduction formulas (2), acts as a "parameter". We think that a better understanding of our topic could be achieved by allowing a change in the parameters envolved, in order to discern the mutual relations between the various relevant properties. For these reasons, we have introduced the space $\mathbb{S}$ and the map $F \mapsto F^*$, since these constructs allow us to change the "parameters" of our phenomena, so to speak, and express the link between the truth values of various statements. Lemma 12 will show that the equivalence between (E) and (W) is a special case of a general equivalence, which has been expressed by introducing the space $\mathbb{S}$ and the map $F \mapsto F^*$, since anything which holds for every $F \in \mathbb{S}$ will also hold for I.

If $F \in \mathbb{S}$, we define $F^\sharp \in \mathbb{S}$ as the unique solution of the second order homogeneous difference equation

$$F_n^\sharp = F_{n-2}^\sharp \cdot F_n \cdot \quad (n \geq 2)$$

with initial data $F_0^\sharp = 1$ and $F_1^\sharp = F_1$. Hence

$$F_0^\sharp = 1, \quad F_1^\sharp = F_1, \quad F_2^\sharp = F_2, \quad F_3^\sharp = F_3 \cdot F_1, \quad F_4^\sharp = F_4 \cdot F_2,$$

Example 10. If $I \in \mathbb{S}$ is the identity sequence, defined in Example 8, then $I_n^\sharp \equiv n!!$

Lemma 11. *If $F \in \mathbb{S}$ and $p, q \in (0, +\infty)$ then*

$$F_{2n}^*(p,q) = p \frac{F_{2n-1}^\sharp}{F_{2n}^\sharp} \quad (n \geq 1) \qquad (6)$$

Proof. Induction. If $n = 1$ then $F_{2n}^*(p,q) = F_2^*(p,q) = F_0^*(p,q)\frac{F_1}{F_2} = p\frac{F_1}{F_2} = p\frac{F_1^\sharp}{F_2^\sharp}$. Assume that (6) holds for n. Then

$$F_{2(n+1)}^*(p,q) = F_{2n+2}^*(p,q)$$
$$= F_{2n}^*(p,q) \cdot \frac{F_{2n+1}}{F_{2n+2}}$$
$$= p\frac{F_{2n-1}^\sharp}{F_{2n}^\sharp} \cdot \frac{F_{2n+1}}{F_{2n+2}} = p\frac{F_{2n+1}^\sharp}{F_{2n+2}^\sharp} \qquad \square$$

Lemma 12. *If $F \in \mathbb{S}$ and $p, q \in (0, +\infty)$ then the following conditions are equivalent:*

$$\frac{F_{2n}^\sharp}{F_{2n-1}^\sharp} \sim \sqrt{p\frac{F_{2n}}{q}} \qquad (W')$$

$$F_{2n}^*(p,q) \sim \sqrt{p\frac{q}{F_{2n}}} \qquad (E')$$

Proof. It suffices to apply (6) $\qquad \square$

Corollary 13. (E) *and* (W) *are equivalent.*

Proof. We specialize Lemma 12 to $F = I$, $p = \pi$, and $q = 2$. Then (W') says that $\frac{(2n)!!}{(2n-1)!!} \sim \sqrt{\pi n}$, and this is (W). On the other hand, since $I^*(\pi, 2) = S$, as seen in (3) and Example 8, then (E') says that $S_{2n} \sim \sqrt{\frac{\pi}{n}}$, and this is (E.2), which is equivalent to (E). $\qquad \square$

3 A proof of (E) and (W)

Since we have established that (E) and (W) are equivalent, in order to prove them it suffices to prove one of them. In Lemma 12 we have seen that the equivalence between (E) and (W) is a special case of a more general equivalence, which has been expressed in terms of the space $\mathbb{S}$ and the map

$$F \mapsto F^*$$

Now, we would like to understand the meaning of the properties (W') and (E') which appear in Lemma 12. Observe that both (W') and (E') contain the initial data (p, q) in their statement. Thus, we would like to understand those properties in terms of the *asymptotic properties* of the sequence F^*, i.e., in terms which do *not* depend on the initial data but only on the behavior of F_n^* for large values of n. We will now show that this may be achieved by using the following notions.

We say that $G \in \mathbb{S}$ has *step 1* if $G_n \sim G_{n+1}$. For example, powers of n and polynomials in n have step 1. If G has step 1 then $G_n \sim G_{n+2}$, but this condition is strictly weaker: We say that a sequence with the latter property has *step 2*. The next result shows that monotonicity acts as a reinforcing condition.

Lemma 14. *If $G \in \mathbb{S}$ has step 2 and it is and monotone then it has step 1.*

Proof. If G is decreasing then

$$\frac{G_n}{G_{n-2}} < \frac{G_n}{G_{n-2}} \cdot \frac{G_{n-2}}{G_{n+1}}$$
$$= \frac{G_n}{G_{n+1}} = \frac{G_n}{G_{n+2}} \cdot \frac{G_{n+2}}{G_{n+1}} < \frac{G_n}{G_{n+2}}$$

hence the result. If G is increasing the inequalities are reversed. $\qquad \square$

If $G \in \mathbb{S}$ has step 1 then, in particular, $G_{2n+1} \sim G_{2n+2}$, but the latter condition is strictly weaker than the former, and we say that a sequence which satisfies the latter condition has *partial step 1*.

Lemma 15. *If $G \in \mathbb{S}$ has step 2 then it has step 1 if and only if it has partial step 1.*

Proof. From $G_{2n+1} \sim G_{2n+3}$ and $G_{2n+1} \sim G_{2n+2}$ we deduce that $G_{2n+2} \sim G_{2n+3}$, which means that $G_{2n} \sim G_{2n+1}$ and since we know that $G_{2n+1} \sim G_{2n+2}$, the conclusion follows. $\qquad \square$

Monotonicity acts as a reinforcing condition also with respect to the notion of having partial step 1, but actually a weaker notion of monotonicity will do. Observe that if $G \in \mathbb{S}$ is monotone then the subsequences $\{G_{2n}\}_n$ and $\{G_{2n+1}\}_n$ are either both increasing or both decreasing, but this condition is strictly weaker than monotonicity, and we call *separately monotone* a sequence which satisfies the latter condition.

Lemma 16. *If $G \in \mathbb{S}$ is separately-monotone then it has step 1 if and only if it has partial step 1.*

Proof. Since the property of having partial step 1 is weaker than the property of having of step 1, it suffices to show that if it has of partial step 1 and separately monotone then it has step 1. The proof of this fact is similar to that of Lemma 14 and will be omitted. □

Now we show that F^* shares some of the properties of F, in a weaker form.

Lemma 17. *If F has step 1 then F^* has step 2. If F is monotone then F^* is separately-monotone.*

Proof. It suffices to observe that (5) may be written as

$$\frac{F_n^*}{F_{n-2}^*} = \frac{F_{n-1}}{F_n} \quad (n \geq 2) \qquad \square$$

Lemma 18. *If $F \in \mathbb{S}$, $p, q \in (0, +\infty)$, and $F_1 = 1$, then the following two conditions are equivalent:*
1. *one of the equivalent conditions of Lemma 12 holds for F, p, and q;*
2. *F^* has partial step 1.*

Proof. The following identity can be shown by induction.

$$F_{2n+1}^*(p, q) = qF_1 \frac{F_{2n}^\sharp}{F_{2n+1}^\sharp} \quad (n \geq 1) \qquad (7)$$

The proof is similar to that of (6) and is omitted. Then (6) and (7) imply that

$$\frac{F_{2n+1}^*(p, q)}{F_{2n+2}^*(p, q)} = \frac{q}{p} F_1 \frac{F_{2n}^\sharp F_{2n+2}^\sharp}{\left(F_{2n+1}^\sharp\right)^2} = \frac{q}{p} F_1 \left[\frac{F_{2n+2}^\sharp}{F_{2n+1}^\sharp}\right]^2 \frac{1}{F_{2n+2}} \qquad (8)$$

If we assume that $F_1 = 1$, then (8) shows that (W') is equivalent to $F_{2n+1}^*(p, q) \sim F_{2n+2}^*(p, q)$. □

Lemma 18 shows that the properties (W') and (E'), appearing in Lemma 12, may be expressed in terms of asymptotic properties of the sequence F^* (and say that F^* has partial step 1). Thus, we now have to show that I^* has partial step 1. Observe that, by Lemma 16, since I^* has step 2, then it has partial step 1 if and only if it has step 1. On the other hand, Lemma 14 shows that, since I^* has step 2, it is sufficient to show that I^* is monotone, in order to show that it has step 1. The fact that I^* is monotone follows from the equality $I^*(\pi, 2) = S$ and from the fact that S is monotone, and this fact in turn follows from its "explicit" representation given in (3).

4   Conclusion

The story about Euler, Stirling, and Wallis is not yet over, we think. The reader has no doubt noticed that the equivalence between (E) and (W) has been obtained through a sequence of intermediate statements, for each of which there is an a priori, direct link with the following one. Is it possible to achieve the equivalence in just one step? Is there an a priori, direct link (or a sequence of these) between (E) and (S) which does not go through (W)? Is it possible to refine the argument involving the map $F \mapsto F^*$?

Bibliography

[1] R. Descartes. *Discourse on Method and The Meditations*. Penguin Books, 1968.

[2] L. Euler. De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt. *Commentarii academiae scientiarum Petropolitanae*, 5:36–57, 1738. [E19] http://eulerarchive.maa.org/pages/E019.html.

[3] L. Euler. Evolutio formulae integralis $\int x^{f-1} dx(lx)^{\frac{m}{n}}$ integratione a valore $x = 0$ ad $x = 1$ extensa. *Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 16:91–139, 1772. [E421] http://eulerarchive.maa.org/pages/E421.html.

[4] D. Gale. The Game of Hex and the Brouwer Fixed-Point Theorem. *The American Mathematical Monthly*, 86(10):818–827, 1979.

[5] C. F. Gauss. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. Friedrich Perthes & I.H. Besser, 1809. https://www.e-rara.ch/zut/doi/10.3931/e-rara-522.

[6] D. B. Herrmann. *The History of Astronomy from Herschel to Hertzsprung*. Cambridge University Press, 1984.

[7] T. P. Jameson. 78.16 The probability integral by volume of revolution. *The Mathematical Gazette*, 78(483):339–340, 1994.

[8] A. N. Kolmogorov. Sur les fonctions harmoniques conjuguées et les séries de Fourier. *Fund. Math.*, 7:24–29, 1925.

[9] S. A. Kripke. *Naming and Necessity*. Basil Blackwell, 1980.

[10] P.-S. Laplace. Mémoire sur la probabilité des causes par les événements. *Mémoires de l'Académie royale des sciences de Paris (Savants étrangers)*, VI:621–656, 1774. Reprinted in Œuvres complètes, 8:27–65, 1891 https://gallica.bnf.fr/ark:/12148/bpt6k77596b/f4.image.

[11] P.-S. Laplace. *Théory analytique de probabilités*. Courcier, 1812. https://gallica.bnf.fr/ark:/12148/btv1b8625611h/f7.image.

[12] P. Levrie and W Daems. Evaluating the probability integral using Wallis's product formula for π. *The American Mathematical Monthly*, 116(6):538–541, 2009.

[13] R. E. Moritz. An elementary proof of Stirling's formula. *Journal of the American Statistical Association*, 23(161):55–57, 1928.

[14] C. P. Nicholas and R. C. Yates. The Probability Integral. *The American Mathematical Monthly*, 57(6):412–413, 1950.

[15] H. L. Royden. *Real Analysis*. Macmillan, second edition, 1968.

[16] E. M. Stein. On limits of sequences of operators. *Annals of Mathematics*, 74(1):140–170, 1961.

[17] J. Wallis. *The Arithmetic of the Infinitesimals*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, 2004.

[18] R. M. Young. 89.36 on evaluating the probability integral: a simple one-variable proof. *The Mathematical Gazette*, 89:252–254, 2005.

[19] A. Zygmund. *Trigonometric Series*. Cambridge Mathematical Library. Cambridge University Press, 2002.

Fausto di Biase [fausto.dibiase@unich.it] is an associate professor of mathematics at the Università "G. D'Annunzio" of Chieti-Pescara. Laurea in Mathematics, Università of Roma "La Sapienza". Ph.D. in Mathematics, Washington University in St. Louis MO USA. His main research interest is the boundary behaviour of harmonic and holomorphic functions.