Proof by Recurrence

Author(s): Sydney Lupton

Source: *The Mathematical Gazette*, Vol. 6, No. 94 (Oct., 1911), pp. 142-148

Published by: [Mathematical Association](#)

Stable URL: <http://www.jstor.org/stable/3603247>

Accessed: 24-02-2016 04:48 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Mathematical Association is collaborating with JSTOR to digitize, preserve and extend access to *The Mathematical Gazette*.

<http://www.jstor.org>

in the first; this one name must, therefore, be A ; and it follows, therefore, that the same name is collated discontinuously in both groups.

(10) Thus we conclude: *In a positive reciprocal collation every conjugate pair separates every other; in a negative reciprocal collation no conjugate pair separates any other; in a zero reciprocal collation, all conjugate pairs have a common unit.*

(11) And: *In a positive reciprocal collation no unit is self-conjugate; in a negative reciprocal collation two units are self-conjugate, and they form a self-conjugate boundary, which separates the two units of every self-conjugate pair; in a zero reciprocal collation it is the same unit in both groups which is collated discontinuously in both groups, and it alone is self-conjugate.*

E. T. DIXON.

(To be continued.)

PROOF BY RECURRENCE.

IN its most simple and general form the proof by recurrence of a formula containing constants and one variable consists in giving a number of successive integral values to the variable, and showing by trial that the formula is true for those values. The next step is to show that the *form* of the result remains the same when $n+1$ is substituted for n , or that each result depends upon the preceding one. If this second step is neglected, as it often used to be, the proof seems to be invalid.

This proof by mathematical or successive induction has been very generally accepted by mathematicians and neglected by logicians.

Todhunter's¹ account of the process and application to the binomial theorem may be accepted as authoritative, from his large acquaintance with previous writers and his special study of the theory of probability: "We prove that if a theorem is true in one case, whatever that case may be, it is true in another case which we may call the *next* case; we prove by trial that the theorem *is* true in a certain case; hence it is true in the next case, and hence in the next to that, and so on; hence it must be true in every case after that with which we began." "Hence if the laws hold when $n-1$ factors are multiplied together, they hold when n factors are multiplied together; but they have been proved to hold when four factors are multiplied together, therefore they hold when five factors are multiplied together, and so on; thus they hold universally."

The article "Induction" (Mathematics) in the *Penny Cyclopædia* by de Morgan has been so long out of print that a mere reference would be of little use: "The method of induction, in the sense in which the word is used in natural philosophy, is not known in pure mathematics. There certainly are instances in which a general proposition is proved by a collection of the demonstrations of different cases, which may remind the investigator of the inductive process, or the collection of the general from the particular. Such instances, however, must not be taken as permanent, for it usually happens that a general demonstration is discovered as soon as attention is turned to the subject.

"There is, however, one particular method of proceeding which is extremely common in mathematical reasoning, and to which we propose to give the name of *successive induction*. It has the main character of induction in physics, because it is really the collection of a general truth from a demonstration which implies the examination of every particular case; but it differs from the process of physics, inasmuch as each case depends upon the one which precedes. Substituting, however, demonstration for observation, the mathematical process bears an analogy to the experimental one,

¹ *Algebra*, 272, 288.

which in our opinion is a sufficient justification of the term 'successive induction.' "There are cases in which the successive induction only brings any term within the general rule when two, three, or more of the terms immediately preceding are brought within it. Thus, in the application of this method to the deduction of the well-known consequence of $x + \frac{1}{x} = 2 \cos \theta$, namely $x^n + \frac{1}{x^n} = 2 \cos n\theta$, it can only be shown that any one case of this theorem is true, by showing that the preceding two cases are true."

The most recent exposition of proof by recurrence is given by Prof. H. Poincaré.¹ It is stated to be the mathematical reasoning *par excellence*, and the general proof of the ordinary rules of arithmetic; further, it is claimed that mathematical induction is an affirmation of a property of the mind itself, and therefore far more certain than induction from observations or experiments which is based upon the belief in the general order of the universe.

Mill,² without discussing the validity of proof by recurrence, remarks: "Such cases as these are but examples of what I have called induction by parity of reasoning, that is not really induction because not involving inference of a general proposition from particular instances."

Stanley Jevons,³ who differs so much from Mill on the general theory of induction, lays down very definitely: "That no finite number of instances can ever prove a general law, or can give us certain knowledge of even one other instance."

Babbage⁴ gives a curious and powerful argument in the same sense from results obtained or obtainable by his calculating machines. He states that a machine can be made which will give a hundred millions and one of the series of natural numbers, but in the following 2761 terms ten thousand times the successive triangular numbers will be automatically added, after which the next 1430 terms follow another law; and then there appears a third new law for 950 terms more. Hence no series of recurrences, however long, can foretell the advent of the term or terms which apparently belong to a different law, but have been provided for in the original design of the machine.

Some of those, who have freely used proof by recurrence, seem to have had lingering doubts as to its validity, even in the case of integral values of the variable, since many mathematicians, such as James Bernoulli, Landen, Epinus, Euler, L'Huilier, Stewart and Lagrange, have given other proofs of the binomial theorem. There seems considerable justification for such doubts, especially when reliance is placed merely upon a series of recurrences proved by trial, and not upon the form of the result in successive instances.

The following formulae give series of 16, 29 and 40 primes:

$$x^2 + x + 17; \quad 2x^2 + 29; \quad x^2 + x + 41;$$

but, of course, all such formulae must fail when the value given to the variable is equal or bears a certain relation to the constant; hence no formula is known which will give a prime for every integral value of the variable. But in the above cases the proof by recurrence might be considered to be ample.

The difference between valid and invalid cases of proof by recurrence is often felt intuitively, but no general method of discriminating seems to be known, nor is it clearly shown why, if it be valid for positive integral values of the variable, it should be invalid in some cases for fractional or negative values.

¹ *Science and Hypothesis*, 9 et seq.

² *Logic*, i. 325.

³ *Principles of Science*, 231.

⁴ *Ninth Bridgewater Treatise*, 34.

A formula may be true and intelligible for all values of the variable except one, thus $\int x^n dx$ is $x^{n+1}/n+1$ except in the case that $n = -1$ when it is $\log x$; or it may be only true and intelligible for one value of the variable; $y - b = \sqrt{-1}(x - a)$ is the equation to a straight line, and is satisfied if $x = a$ and $y = b$, which represents a point in the plane of reference, but every other point of the line is in the plane perpendicular to the plane of the paper.

More generally a formula is true and intelligible for certain values of the variable, but untrue or unintelligible for intermediate values. When these singular values are expressed graphically, they appear as conjugate points or pointed branches.

For positive values of x the equation $y = ax^2 + \sqrt{x} \sin bx$ gives a continuous curve looped about a parabola, which it cuts when bx is a multiple of π . But when x is negative y is impossible except when $\sin bx = 0$, or the figure consists of an infinite number of conjugate points situated on the negative branch of the parabola with abscissae $\pi/b, 2\pi/b, 3\pi/b$, etc.

Such cases suggest reasons for excluding negative but not fractional values of the variable.

Proof by recurrence may also be regarded as a special case of the probability of causes, or of the probability of future events derived from experience. The question to be decided is, a given formula gives true results in certain cases, what is the chance that the formula is exact and will give true results in every case, or that it is inexact and will fail in certain cases?

Problems of this kind are generally considered to be identical with the determination by a succession of drawings of the relative numbers of white and black balls in an urn, the total number of which may be known, unknown, or infinite.

The chance of drawing a white ball evidently depends upon the relative number of white and black balls in the urn. If the number of each kind be very large compared to the number of drawings, or if each drawn ball be replaced, this relation remains constant; but it more nearly corresponds to the case of mathematical or physical formulae, if the drawn balls are not replaced and the relation is treated as variable.

Before a trial is made, we may make any *assumption* as to the relative number of white and black balls in the urn. The most obvious supposition is that every possible combination is equally possible, or that the probability of each is $\frac{1}{n+1}$. Or we may suppose that the white and black balls have

been selected by tossing a coin, when the probability of each combination is expressed by a term of the binomial expansion $(\frac{1}{2} + \frac{1}{2})^n$ or the relative probability is $1, \frac{n}{1}, \frac{n \cdot n-1}{1 \cdot 2}, \frac{n \cdot n-1 \cdot n-2}{1 \cdot 2 \cdot 3}$, etc.

It seems likely that the greater the number of trials, the greater is the chance of estimating the relative numbers of white and black balls remaining in the urn. But to take an extreme case, suppose that $n-1$ white balls have been drawn and not replaced, leaving only one in the urn, what is the probability that the last ball is white?

There are only two possible cases: the urn may have originally contained n white balls, or it may have contained $n-1$ white balls and one black. In the first case the remaining ball is certainly white, and in the second certainly black. On the *assumption* that each case is, a priori, equally probable, the chance that the last ball is white is $\frac{1}{2}$. On the second a priori *assumption* the possible hypotheses give for the probability of the event

in the first case 1, in the second case $\frac{\frac{n-1}{n}}{\frac{1}{n}}$ or $\frac{1}{n}$, and in the other cases 0;

hence the probability that the fifth ball is white is

$$\frac{1}{1+n \cdot \frac{1}{n} + 0} = \frac{1}{2}, \text{ as before.}$$

Suppose, as before, that the unknown proportion of white balls, which is identical with the probability x of drawing a white ball, represents the chance that a formula is valid or universally true, and that the chance of drawing a black ball $(1-x)$ is the chance that the formula is not universally true, or that the favourable instances have occurred by accident, and to obviate arithmetical difficulties let each drawn ball be replaced and the urn shaken.

The accepted treatment of the problem is somewhat as follows: since the relative proportion of white and black balls is unknown, assume that all values of x are equally probable. If $m+n$ drawings give m white and n black balls, the probability which any assumed value of x gives for the observed event is $x^m(1-x)^n$. On differentiating this expression, the maximum value for x is found to be $m/m+n$.

Bayes assumed that x may vary infinitesimally between the limits 1 when all the balls are white, and 0 when all the balls are black. But if a white occurs more frequently than a black ball, it is a fair assumption that the probability of a cause producing a white ball is $> \frac{1}{2}$, and increases with the number of occurrences. Hence the probability of a cause which produces a white ball is

$$P = \frac{\int_{\frac{1}{2}}^1 x^m(1-x)^n dx}{\int_0^1 x^m(1-x)^n dx}.$$

But a black ball at once shows that the formula is invalid, and if $n=0$ the probability is $P = \int_{\frac{1}{2}}^1 x^m dx / \int_0^1 x^m dx$ or $1 - \left(\frac{1}{2}\right)^{m+1}$.

Comparatively few trials give high values for P , if $m=5$, $P=63/64$. But it must be clearly remembered that this only means that there is a high probability, that the chance of the existence of a cause lies between $\frac{1}{2}$ and 1, or that the formula is more likely to be valid than invalid. The same formula may be used to calculate the probability that the existence of a cause is greater than any assigned limit. After 100 white balls have been drawn, to find the probability that the urn contains at least 9999 white to one black ball. $P = 1 - \left(\frac{9999}{10000}\right)^{101} = 1/100$ nearly.

The probability of a considerable chance that the formula is universally true is small, and to increase it to $\frac{1}{2}$ about 6930 trials would be required. According to Bayes' theorem, then, no finite number of recurrences can make a formula more than highly probable, and even to get so far, the successful trials must be very numerous.

Assuming the same data, Laplace obtained a different solution. The probability that x lies between x and $(x+dx)$ is expressed by $cx^m(1-x)^n dx$, where c is a constant independent of x . But the probability that x lies between 0 and 1 is certain.

$$\therefore 1 = c \int_0^1 x^m(1-x)^n dx = cB.(m+1, n+1) = c \frac{\Gamma(m+1)\Gamma(n+1)}{\Gamma(m+n+2)};$$

$$\therefore c = \frac{m+n+1}{m!n!}.$$

Hence the probability of the hypothesis, that the ratio between the white and black balls is $x/(1-x)$, is $\frac{m+n+1}{m!n!} x^m(1-x)^n dx$. And on the assumption that this ratio is correct if $(\mu+\nu)$ future drawings be made, the most probable

case is that the ratio μ/ν is the same as the ratio m/n , and the probability of this occurrence is expressed by $\frac{\mu+\nu}{\mu} x^\mu (1-x)^\nu dx$. But the probability of the compound event is the product of the two separate probabilities, or

$$P = \frac{\mu+\nu}{\mu} \frac{m+n+1}{\nu} \int_0^1 x^{\mu+1} (1-x)^{n+\nu} dx \\ = \frac{\mu+\nu}{\mu} \frac{m+n+1}{\nu} \frac{m+\mu}{m+\mu+n+\nu+1}.$$

This formula has been frequently applied by Laplace and others generally to find the chance of one future occurrence after a series of favourable trials, in which case $n=0$, $\mu=1$, $\nu=0$, and the factorials reduce to $P = \frac{m+1}{m+2}$.

Thus Laplace finds a very high probability that the sun will rise to-morrow. After 6930 white balls have been drawn, the chance that the next one will be black is only 1/6932, which may be compared with the somewhat indefinite result obtained by Bayes' theorem.

Bertrand,¹ after giving the proof of the formula, significantly remarks: "The applications which have been made of this formula are almost all without foundation." He derides the exactness of the result obtained in the case of the sun.

Boole² dissents from the validity of both these solutions: "When the defect of data is supplied by hypothesis, the solutions will in general vary with the nature of the hypothesis assumed; so that the question still remains only more definite in form, whether the principles of the theory of probabilities serve to guide us in the election of such hypothesis. I have already expressed my conviction that they do not."

If the validity of Laplace's formula be admitted, no finite number of favourable instances can make the validity of a general formula certain, and a very long series of recurrences is required to render comparatively few future terms even somewhat highly probable.

Whether mathematical induction or proof by recurrence is to be regarded as a true induction or not, depends upon the exact definition given of the latter process.

Logicians differ much as to this definition, and some have even given two or more definitions not exactly in accordance with one another.

Whately³ defines induction as "A kind of argument which infers respecting a whole class what has been ascertained respecting one or more individuals of that class."

Mill⁴ states, "Induction may be defined, the operation of discovering and proving general propositions"; where 'prove' must be taken in its original sense of testing by trial or rendering probable. And again, "Induction is that operation of the mind, by which we infer that, what we know to be true in a particular case or cases, will be true in all cases which resemble the former in certain assignable respects."

It seems as though proof by recurrence falls within any of these definitions, but Mill specially classes it among inductions improperly so-called, and instances Newton's proof of the binomial theorem.

The steps by which a valid induction is reached appear to be four in number:

We collect a considerable number of similar instances, or, in the words of Whewell, we form a colligation of facts.

¹ *Calcul des Probabilités*, 172.

² *Laws of Thought*, xx.

³ *Logic*, 267.

⁴ *Logic*, III. i. and ii.

We abstract in each case the circumstances common to all and material to our argument from others which are not common to all or immaterial to our argument.

We then connect the common circumstances by a rule or formula which contains them all. In attempting to find a rule or formula it is often difficult to distinguish and isolate the true argument which connects the instances. When this has been effected, the form of the formula may often be obtained from theoretical considerations, or from inspection of the graphical representation of the instances, or finally, in case of failure, recourse must be had to the more or less blind trial of hypotheses.

Lastly, the rule or formula must be verified by showing that it applies to other facts either similar to or different from those already ascertained.

The colligation of facts is a necessary preliminary to every induction, since a single instance is a very insecure foundation, but it is not itself an induction. The idea of induction contains as a necessary element a leap into the unknown, or an assumption that cases which have not been considered will fall under the same rule as those which have been already investigated. This leap of the mind into the unknown introduces an element of doubt into every induction, which is more or less removed by subsequent verification. If the verification is perfect by the examination of every possible instance, the doubt ceases, but the rule also ceases to be an induction and becomes a mere colligation of facts.

The above view of induction may be applied to Newton's¹ statement of his discovery of the binomial theorem. "I applied myself, therefore, to discover a method by which the first two terms of this series might be derived from the rest; and I found that if for the second figure or numerical term I put m , the rest of the terms would be produced by the continual multiplication of the terms of the series: $\frac{m-0}{1} \times \frac{m-1}{2} \times \frac{m-2}{3} \times \frac{m-3}{4}$, etc.

For instance, if the second term be put for 4, there will arise $4 \times \frac{m-1}{2}$, that is 6, which is the third term; the fourth term will be $6 \times \frac{m-2}{3}$, that is 4; the fifth term will be $4 \times \frac{m-3}{4}$, that is 1; and the sixth term will be $4 \times \frac{m-4}{5}$, that is 0, which shows that the series is here terminated in this case.

"This being found, I applied it as a rule to interpolate the above-mentioned series."

It seems then that Newton was led to the form of the coefficients by the mere consideration of successive terms in one or more expansions. From the consideration of these instances his mind leapt to the dependence of each term on the one before it, and the rule connecting them. His intuitive genius caused him to be satisfied with a comparatively small amount of verification, and extended the theorem in the same manner to negative and fractional indices. Universal experience has since shown his assumption to be correct, and various general proofs have been given.

If a mathematical formula is true, it is generally assumed to be true universally and exactly, except in those comparatively rare instances in which small terms have been omitted or approximate values introduced for some definite purpose. Hence a mathematician has no compunction in applying a formula to cases intermediate, between, or far beyond those which have been rendered certain by trial.

Experimental values, on the other hand, are not exactly true, and can only be more or less approximately represented by a formula. Hence,

¹ Horsley, iv. 540.

unless his experimental values are fairly close together, a physicist is chary in applying his formulae to intermediate values, or in any case much beyond the extreme values proved by his experiments.

In several well-known cases, *e.g.* Fermat, too much reliance seems to have been placed on proof by recurrence; and the consideration of a simple case may be suggestive.

The formula n^2+n+41 gives a prime for all values of n from 0 to 39; the following 81 values give primes in all except 18 cases, and in those cases the factors are two primes. Trusting entirely to proof by recurrence, the first thirty-nine values of n would seem to show that the formula always gives a prime.

Excluding 0, the chance that a value for $n < 121$ gives a prime is $\frac{102}{120}$, and if the first trial gives a prime, the chance of a prime at the second trial is $\frac{101}{119}$, and so on. Hence, writing x for one less than the number of trials, the chance of prime for $x+1$ trials is

$$\frac{102}{120} \cdot \frac{101}{119} \cdot \frac{100}{118} \cdots \frac{102-x}{120-x} \text{ or } \frac{102}{120} \times \frac{120-x}{102-x}.$$

And the chance of not-prime appearing in $x+1$ trials is

$$1 - \frac{102}{120} \times \frac{120-x}{102-x}.$$

When any fractional value is given to this expression, the value of x can be found by trial.

Thus, how many trials must be made that the chance of falling upon not-prime is $\frac{9999}{10000}$?

The value of x is found to be about 44, and the chance that forty-five trials give white is $\frac{1}{11360}$.

If $x=102$, $\frac{102}{120} \times \frac{18}{0} = 1/1.0867 \times 10^{21}$, or the chance that a negative in-

stance is obtained is very great. If 104 or more trials are made, a negative instance must be obtained, or the value of the factorial expression should be 0. Since $[-1 = \Gamma 0 = \infty$, which is also true for all negative integral values of the gamma function, this is seen to be the case. SYDNEY LUPTON.

APPRECIATIVE REMARKS ON THE THEORY OF GROUPS.¹

WHILE it is clearly impossible for the average high school teacher of mathematics to become familiar with all the modern branches of this subject, it is desirable that he should not be totally ignorant of any extensive branch. The views of a number of eminent mathematicians often furnish one of the simplest as well as one of the most reliable introductions to the nature and the scope of a difficult subject. The following list of quotations has been prepared for the purpose of providing such an introduction for one subject. The list could easily be extended, but the variety and the standing of the mathematicians quoted are perhaps sufficient to inspire confidence. The

¹ Reprinted by permission from *School Science and Mathematics*, with some alterations and additions by the author.