WILEY-VCH

Colm T. Whelan

A First Course in Mathematical Physics

Colm T. Whelan

A first Course in Mathematical Physics

Related Titles

Halliday, D., Resnick, R., Walker, J.

Principles of Physics, Ninth Edition, Volume 1 (Chapter 1-20) International Student Version

9th Edition
2011
Print ISBN: 978-0-470-56837-8

Boas, M.L.

Mathematical Methods 3E

3rd Edition
2010
Print ISBN: 978-0-470-94796-8

Kusse, B., Westwig, E.A.

Mathematical Physics
Applied Mathematics for Scientists and Engineers

1998
Print ISBN: 978-0-471-15431-0

Courant, R., Hilbert, D.

Methods of Mathematical Physics
Volume 2

1989
Print ISBN: 978-0-471-50439-9
ISBN: 978-3-527-61723-4
Adobe PDF ISBN: 978-3-527-61724-1

Fanchi, J.R.

Math Refresher for Scientists and Engineers, Third Edition

3rd Edition
2006
Print ISBN: 978-0-471-75715-3
Mobi-lite ISBN: 978-0-470-32623-7
Adobe PDF ISBN: 978-0-471-79154-6
ISBN: 978-0-471-79155-3

Understanding Physics

1998
Print ISBN: 978-0-471-97553-3

Vaughn, M.T.

Introduction to Mathematical Physics

2007
Print ISBN: 978-3-527-40627-2
ISBN: 978-3-527-61885-9
Adobe PDF ISBN: 978-3-527-61886-6

Fujita, S., Godoy, S.V.

Mathematical Physics

2009
Print ISBN: 978-3-527-40808-5

Masujima, M.

Applied Mathematical Methods in Theoretical Physics

2nd Edition

2009

Print ISBN: 978-3-527-40936-5

ISBN: 978-3-527-62774-5

Adobe PDF ISBN: 978-3-527-62775-2

Nonequilibrium Statistical Physics

2013

Print ISBN: 978-3-527-41092-7

Adobe PDF ISBN: 978-3-527-67057-4

ISBN: 978-3-527-67139-7

Toral, R., Colet, P.

Stochastic Numerical Methods

An Introduction for Students and Scientists

2014

Print ISBN: 978-3-527-41149-8

MobiPocket ISBN: 978-3-527-68311-6

ePub ISBN: 978-3-527-68312-3

Adobe PDF ISBN: 978-3-527-68313-0

WOL obook PDF ISBN: 978-3-527-68314-7

Martin, B.R., Shaw, G.

Mathematics for Physicists

2015

Print ISBN: 978-0-470-66022-5

ePub ISBN: 978-1-118-67661-5

Adobe PDF ISBN: 978-1-118-67663-9

Colm T. Whelan

A first Course in Mathematical Physics

WILEY-VCH
Verlag GmbH & Co. KGaA

Author

Professor Colm T. Whelan
Old Dominion University
Department of Physics
4600 Elkhorn Avenue
VA
United States

Cover
math formula: "fotolia_© shotsstudio";
sepia-Motive: "fotolia_© agsandrew"

Library of Congress Card No.: applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.d-nb.de>.

Print ISBN: 978-3-527-41333-1
ePDF ISBN: 978-3-527-68713-8
ePub ISBN: 978-3-527-68715-2
Mobi ISBN: 978-3-527-68714-5

Cover Design Adam-Design, Weinheim, Germany
Typesetting SPi Global, Chennai, India
Printing and Binding

Printed on acid-free paper

For Agnieszka, of course

Contents

Preface

This book is an elementary introduction to the mathematics needed for students taking undergraduate classes in the physical sciences. My ambition is to present in a simple and easily intelligible way the core material they will need for their courses and help them to uncover the character of the physical laws, which can sometimes, unfortunately, be obscured by a lack of understanding of and sympathy with the precise mathematical language in which, perforce, they have to be expressed. I have emphasized the direct connection between the conceptual basis of the physics the students are about to learn and the mathematics they are studying. The first part of the book introduces the core mathematics and while I have included numerous applications in this section I felt that a fuller understanding could only be achieved if the physical context was given a more complete introduction. In the second part of the book, I have given a series of brief overviews of some of the more beautiful and conceptually stimulating areas of physics.

The material in this book is designed for a one- or two-semester course to be ideally taken at the start of the sophomore year. I assume next to no previous knowledge. The first two chapters give a brief overview of some of the more elementary mathematics one might hope that students bring with them from high school and their first year. The choice of material included in the course involved some painful pruning. I am all too aware of the numerous shortcuts I have taken, the many important interesting and beautiful diversions I have ignored, and the rigor I have forsaken all with one principle object in view to equip the students with a basic mathematical toolkit that will allow them to enjoy and profit from their higher level physics courses. I rather hope that the student will be sufficiently intrigued by the exposure to the interesting topics covered in this text to seek out some other more formal and rigorous specialist mathematics courses. A large number of problems have been included, for none of which is the use of a calculator needed. These form an integral part of the book and students are strongly encouraged to attempt all of them. Instructors can get a complete solution manual from the publisher (http://www.wiley-vch.de/supplements/).

Norfolk, September 2015 *C T W*

Part I
Mathematics

1
Functions of One Variable

1.1
Limits

It is often said that most mathematical errors, which get published, follow the word "clearly" and involve the improper interchange of two limits. In simple terms, a "limit" is the number that a function or sequence "approaches" as the input or index approaches some value. For example, we will say that the sequence $x_n = \frac{1}{n}$ approaches the limit 0 as n moves to infinity Or, in other words, we can make x_n arbitrarily small by choosing n big enough. We often write this as

$$\lim_{n\to\infty} x_n = 0$$

We can also take the limit of a function, for example, if $f(x) = x^2$ then

$$\lim_{x\to 2} f(x) = 4$$

A sequence of numbers x_n is said to converge to a limit x if we can make the difference $|x - x_n|$ arbitrarily small by making n big enough. If such a limit point does not exist, then we state that the sequence diverges. For example, the sequence of integers

$$x_n = 1, 2, 3, \ldots$$

is unbounded as $n \to \infty$, while the sequence

$$x_n = 1 + (-1)^n$$

oscillates and never settles down to a limit. More formally, we state

Definition 1.1. *Let f be a function defined on a real interval I then the limit as $x \to a$ exists if there exists a number l such that given a number $\epsilon > 0$ no matter how small, we can find a number $\delta > 0$, where for all $x \in I$ satisfying*

$$|x - a| < \delta$$

we have

$$|f(x) - l| < \epsilon$$

A first Course in Mathematical Physics, First Edition. Colm T Whelan.

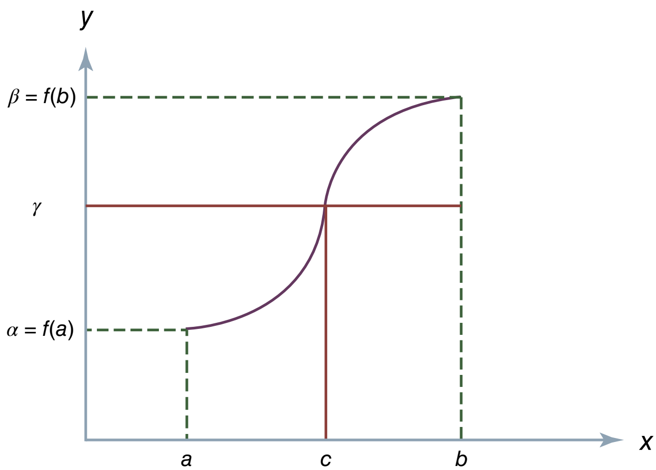

Figure 1.1 If $y = f(x)$ is a continuous function on $[a, b]$ if we pick any value, γ, that is between the value of $\alpha = f(a)$ and the value of $\beta = f(b)$ and draw a line straight out from this point, the line will hit the graph in at least one point with an x value between a and b.

Notice that we do not necessarily let x ever reach a but only get infinitesimally close to it. If in fact $f(a) = l$, then we state that the function is continuous. Intuitively, a function that is continuous on some interval $[a, b]$ will take on all values between $f(a)$ and $f(b)$ (Figure 1.1). For a more formal discussion see [1]. An intuitively obvious result is the intermediate value theorem.

Theorem 1.1. *Let f be a continuous function on a closed interval $[a, b]$ $\alpha = f(a), \beta = f(b)$. If γ is a number such that $\alpha < \gamma < \beta$, then there exists a number c such that*

$$f(c) = \gamma$$

Proof: For a formal proof see for example [1] ■

Consider the sequence of partial sums

$$S_n = \sum_{j=1}^{n} x_n \tag{1.1}$$

if the sequence of partial sums converges to some limit S as $n \to \infty$ then we say that the infinite series $\sum_{j=0}^{\infty} x_j$ is convergent.

Example 1.1. *The geometric series*

Let

$$s_n = \sum_{j=1}^{n} ax^j$$

then

$$\begin{aligned}(1-x)s_n &= \sum_{j=1}^{n} ax^j - ax^{j+1}\\ &= a - ax + ax - ax^2 + ax^2 + \cdots + ax^{n-1} - ax^n + ax^n - ax^{n+1}\\ &= a(1-x^{n+1})\end{aligned}$$

hence

$$s_n = \frac{a(1-x^{n+1})}{1-x}$$

Clearly, therefore, if $|x| < 1$ *the series converges. Its value being given by*

$$\sum_{j=1}^{\infty} ax^j = \frac{a}{1-x}$$

if $|x| \geq 1$, *the series diverges.*

1.2 Elementary Calculus

Assume that we are observing an object moving in one dimension. We measure its position to be x_0 at time $t = t_0$ and $x_0 + \Delta x$ at time $t = t_0 + \Delta t$, thus its average speed is

$$\bar{v} = \frac{(x_0 + \Delta x) - x_0}{(t_0 + \Delta t) - t_0} = \frac{\Delta x}{\Delta t} \tag{1.2}$$

Of course, this is only an average value the object could accelerate and decelerate during the time interval; if we need to know its speed at any given point, then we must shorten the time interval, and to know the "instantaneous" speed at a time $t = t_0$, we need to let Δt lead to zero, that is,

$$v(t) = \lim_{\Delta t \to 0} \frac{x_0 + \Delta x - x_0}{t_0 + \Delta t - t_0} \tag{1.3}$$

This motivates us to define the derivative of a function.

Definition 1.2. *If f is only a function of x, then the first derivative of f at x is defined to be*

$$\frac{df(x)}{dx} \equiv \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{1.4}$$

If this limit exists, then the function is said to be differentiable. The function f is said to be continuously differentiable if the derivative $f'(x)$ exists and is itself a continuous function.

Frequently, we use the notation $f'(a)$ as a shorthand, that is,

$$f'(a) = \frac{df(x)}{dx}|_{x=a}$$

Example 1.2. *If $f(x) = x$, then*

$$f'(x) = \frac{df(x)}{dx} = \lim_{h\to 0} \frac{x+h-x}{h} = 1$$

If $f(x) = c$, where c is a constant, then $f(x+h) - f(x) = 0$ for all h, consequently $f'(x) = 0$. A partial converse to this result is as follows. If $f('x) = 0$ on some interval I, then $f(x) = c$ on I, where c is a constant. This is a consequence of the intermediate value theorem; see Problem 1.1. Clearly if

$$f'(x) = g'(x) \text{ on } I \tag{1.5}$$

then

$$f(x) = g(x) + c \text{ on } I \tag{1.6}$$

where c is a constant.

1.2.1 Differentiation Products and Quotients

Assuming $f(x)$ can be written as the product of two functions, for example,

$$f(x) = u(x)v(x)$$

then

$$\frac{df(x)}{dx} = \lim_{h\to 0} \frac{u(x+h)v(x+h) - u(x)v(x)}{h} \tag{1.7}$$

We may rewrite the numerator in (1.7) as

$$u(x+h)[v(x+h) - v(x)] + v(x)[u(x+h) - u(x)]$$

in the limit as $h \to 0$ $u(x+h) \to u(x)$ and

$$\lim_{h\to 0} \frac{u(x+h) - u(x)}{h} = \frac{du(x)}{dx}$$

$$\lim_{h\to 0} \frac{v(x+h) - v(x)}{h} = \frac{dv(x)}{dx}$$

it immediately follows that

$$\frac{d(u(x)v(x))}{dx} = u(x)\frac{dv(x)}{dx} + v(x)\frac{du(x)}{dx} \tag{1.8}$$

It is also possible to show, Problem 1.10, that

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \tag{1.9}$$

Lemma 1.1.

$$\frac{dx^N}{dx} = Nx^{N-1}$$

Proof: If $N = 1$, clearly true since $\frac{dx}{dx} = 1$; assume to be true for all $N \leq N_0$, consider

$$f(x) = x^{N_0+1} = xx^{N_0}$$

then from (1.8)

$$\frac{df(x)}{dx} = x^{N_0}\frac{dx}{dx} + x\frac{dx^{N_0}}{dx} \tag{1.10}$$

Now, by assumption

$$\frac{dx^{N_0}}{dx} = N_0 x^{N_0-1}$$

and as we have already seen

$$\frac{dx}{dx} = 1$$

hence

$$\frac{df(x)}{dx} = x^{N_0} + xN_0x^{N_0-1} = x^{N_0}[N_0 + 1] \tag{1.11}$$

Hence, by principle of induction, true for all integers. ∎

1.2.2 Chain Rule

Assume that

$$f(x) = u(v(x))$$

For example,

$$\begin{aligned} v(x) &= x^2 \\ u(y) &= \sqrt{1-y} \\ \Rightarrow f(x) &= \sqrt{1-x^2} \end{aligned}$$

then for such a function:

Lemma 1.2. *If v is differentiable at the point x and that u is differentiable at the point $y = v(x)$, then*

$$\frac{df(x)}{dx} = u'(v(x))v'(x) \tag{1.12}$$

or in other words

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}$$

Proof:

$$\frac{u(v(x+h)-u(v(x))}{h} = \frac{u(v(x+h)-u(v(x))}{v(x+h)-v(x)}\frac{v(x+h)-v(x)}{h} \tag{1.13}$$

We can now take the limit as $h \to 0$ and we have the result. In fact, to be more rigorous, we should worry about the possibility of $v(x+h)-v(x)$ passing through 0. For a treatment where this problem is explicitly dealt with see [2]. ■

Example 1.3. *Newton's second law can be written as*

$$\begin{aligned} F = ma &= m\frac{dv}{dt} \\ &= m\frac{dx}{dt}\frac{dv}{dx} = mv\frac{dv}{dx} \\ &= \frac{d[\frac{1}{2}mv^2]}{dx} \end{aligned} \tag{1.14}$$

Thus, force can be defined as the rate of mass times acceleration or as the rate of change of the kinetic energy with distance

1.2.3
Inverse Functions

Consider the functions shown in Figure 1.2. Both are continuous but for $f_2(x)$ the equation

$$y = f_2(x)$$

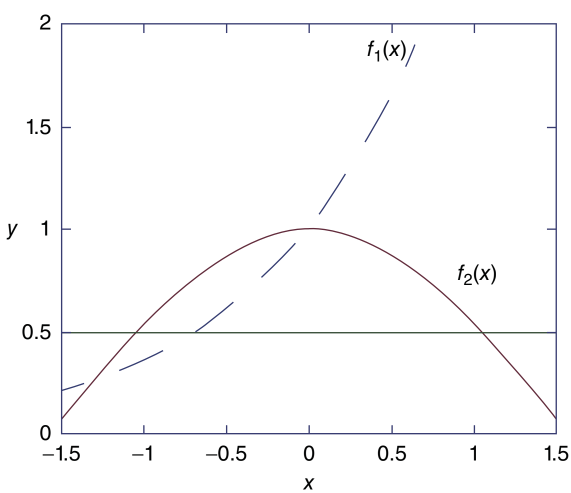

Figure 1.2 Over the range shown, the function $f_1(x)$ is invertible, $f_2(x)$ is not.

does not have a unique solution for $y = \frac{1}{2}$, while the equivalent equation for f_1 will have such a solution. The difference between the two functions is that f_1 is strictly increasing over the entire interval but f_2 is not.

Definition 1.3. *If f is a continuous strictly increasing function on $[a, b]$, with $\alpha = f(a), \beta = f(b)$, then from the intermediate value theorem, Theorem 1.1, we know that the set*

$$\{y = f(x) | a \leq x \leq b\}$$

forms the interval $[\alpha, \beta]$. We may define a function g

$$\begin{aligned} g : [\alpha, \beta] &\to [a, b] \\ g(f(x)) &= \quad x \\ f(g(y)) &= \quad y \end{aligned}$$

It is clear that we could just as well have constructed an inverse for a strictly decreasing function. The only time we will have a problem is when $x_1 \neq x_2$ but $f(x_1) = f(x_2)$. Usually, we write g as f^{-1}. Even for f_2, we can define an inverse if we agree to only look at the intervals $-1.5 \leq x < 0$ and $0 < x \leq 1.5$ separately. We can thus talk about a local inverse, that is, given a point x_0 if we can find an interval around it for which the function f is strictly increasing or decreasing, then we can find an inverse valid in this region. We know that (see Problem 1.3) a function is strictly increasing or decreasing on an interval once its derivative does not change sign; so if the derivative is continuous and we are not at a point x_0, where $f'(x_0) = 0$, then we can always find an interval, however, small so that the function is locally invertible. More formally, we may state the inverse function theorem.

Theorem 1.2. *For functions of a single real variable, if f is a continuously differentiable function with nonzero derivative at the point x_0, then f is invertible in a neighborhood of x_0, the inverse is continuously differentiable and*

$$\frac{df^{-1}(y)}{dy} = \frac{1}{f'(x)}$$

where $x = f^{-1}(y)$.

Proof: If f has a nonzero derivative at x_0, then it follows that there is a interval around x_0 where it is either increasing or decreasing then (Problem 1.4), f^{-1} is continuous. Let $\alpha < y_0 < \beta; y_0 = f(x_0), y = f(x)$, then

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x)-f(x_0)}{x-x_0}}$$

since f^{-1} is continuous the result follows. ∎

1.3 Integration

There are number of equivalent ways of looking at integrals. Perhaps the most intuitive is to consider

$$I = \int_a^b f(x)dx$$

as the area in the plane bounded by the curves $y = f(x), y = 0, y = f(a), y = f(b)$. Conventionally, we often describe this quantity as the area under the curve $y = f(x)$; see Figure 1.3. As a first approximation, we could simply assume that the function $y = f(x)$ could be approximated by its initial value $y = f(a)$ over the entire range, and we would have

$$I \approx f(a)(b - a)$$

See Figure 1.3(a). Now, we clearly lose some area by this approximation. We can improve it by taking a point c, with $a \leq c \leq b$ and approximating the integral by two rectangles of area $f(a)(c - a)$ and $f(b)(b - c)$.

We can continue this process by adding more and more subintervals. If

$$a = \chi_0 < \chi_1 < \chi_2 < \cdots < \chi_n = b \tag{1.15}$$

then we can approximate

$$I \approx \sum_{j=1}^{n} f(x_i)(\chi_i - \chi_{i-1}) \tag{1.16}$$

where

$$\chi_{i-1} \leq x_i \leq \chi_i$$

If we make these intervals arbitrarily small, that is, let $n \to \infty$, then we should get an accurate measure of the area under the curve. This prompts the following definition.

Definition 1.4. *The integral from a to b of the function f is given by*

$$I = \int_a^b f(x) = \lim_{n\to\infty} \sum_{j=1}^{n} f(x_i)(\chi_i - \chi_{i-1}) \tag{1.17}$$

where

$$a = \chi_0 < \chi_1 < \chi_2 < \cdots < \chi_{i-1} \leq x_i \leq \chi_i < \cdots < \chi_n = b$$

We can use the intermediate value theorem to establish the following theorem.

Theorem 1.3. *Mean value theorem for integrals. Let f be continuous on* $[a, b]$, *then there exists a* $c \in (a, b)$ *s.t.*

$$\int_a^b f(x)dx = (b - a)f(c)$$

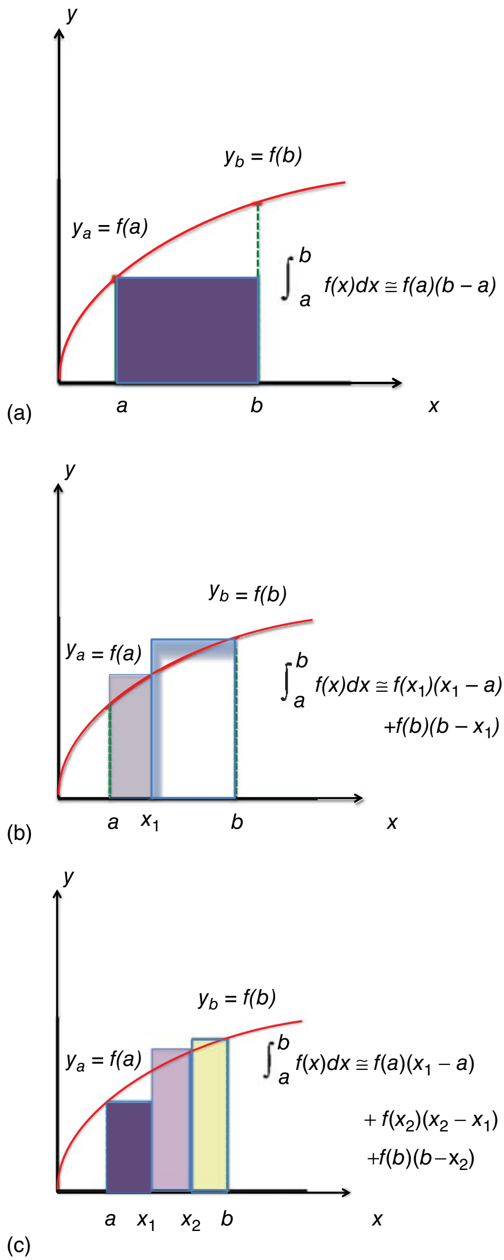

Figure 1.3 (a) Approximating the value of the integral as $f(a)(b-a)$, (b) picking a point x_1 where $a \leq x_1 \leq b$ and approximating integral as $\int_a^b f(x)dx \approx f(x_1)(x_1 - a) + f(x_1)(b - x_1)$, and (c) picking another point x_2 where $x_1 \leq x_1 \leq b$ and approximating integral as $\int_a^b f(x)dx \approx f(x_1)(x_1 - a) + f(x_2)(x_2 - x_1) + f(b)(b - x_2)$.

Proof: From Definition 1.2, if m is the minimum value of f on $[a, b]$ and M its maximum, then

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$$

Hence, by the intermediate value theorem, there exists $c \epsilon [a, b]$ s.t.

$$f(c) = \frac{\int_a^b f(x)dx}{b-a}$$

■

We can define a function

$$F(x) = \int_a^x f(y)dy \tag{1.18}$$

Theorem 1.4.

$$\frac{dF(x)}{dx} = f(x)$$

Proof:

$$F(x+h) = \int_a^{x+h} f(y)dy$$
$$F(x) = \int_a^x f(y)dy$$

$$F(x+h) - F(x) = \int_x^{x+h} f(y)dy \tag{1.19}$$

Now, if h is sufficiently small, we can take $f(y) = f(x)$ over the entire interval and

$$\frac{F(x+h) - F(x)}{h} = \frac{x+h-x}{h} f(x) = f(x) \tag{1.20}$$

take the limit $h \to 0$ and the result follows immediately. ■

We note that the constant a is entirely arbitrary. Theorem 1.4 is rather grandly known as the fundamental theorem of calculus, and it essentially states that integration is the inverse process to differentiation. It has an important corollary.

Corollary 1.1. *Assume that g is continuously differentiable function that maps the real interval* $[a, b]$ *onto the real interval I and that f is a continuous function that maps I into* $\mathbb{R}$. *Then,*

$$\int_{g(a)}^{g(b)} f(x) = \int_a^b f(g(t))g'(t)dt$$

Proof: The function $f(g(t))g'(t)$ is continuous on $[a, b]$ just as f is; therefore, both $\int_{g(a)}^{g(b)} f(x)dx$ and $\int_a^b f(g(t))g'(t)dt$ exist. All we have to show is that they are, in