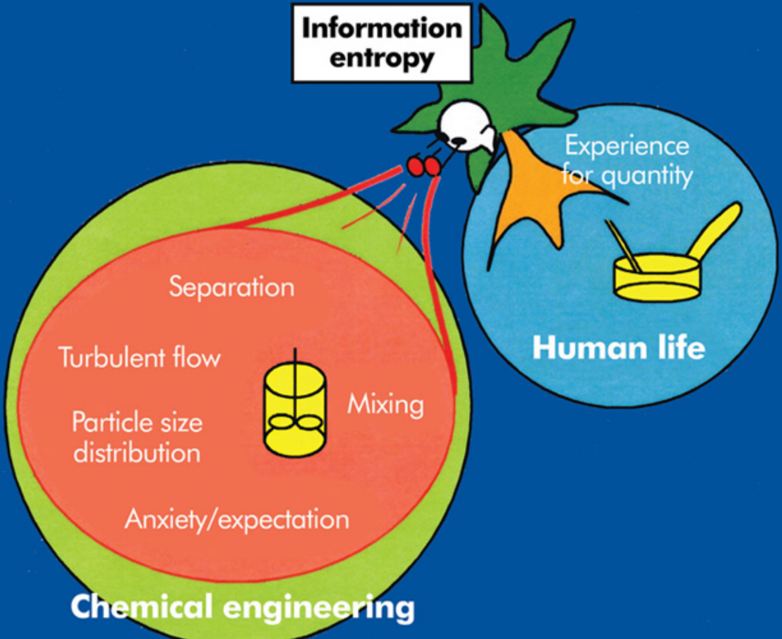
KOHEI OGAWA

CHEMICAL ENGINEERING

A NEW PERSPECTIVE



CHEMICAL ENGINEERING

A NEW PERSPECTIVE

CHEMICAL ENGINEERING

A NEW PERSPECTIVE

Kohei Ogawa
Tokyo Institute of Technology, Japan

ELSEVIER

Amsterdam • Boston • Heidelberg • London • New York • Oxford • Paris
San Diego • San Francisco • Singapore • Sydney • Tokyo

Elsevier
Radarweg 29, PO Box 211, 1000 AE Amsterdam, The Netherlands
Linacre House, Jordan Hill, Oxford OX2 8DP, UK

First edition 2007

Copyright © 2007 Elsevier B.V. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system
or transmitted in any form or by any means electronic, mechanical, photocopying,
recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights
Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333;
email: permissions@elsevier.com. Alternatively you can submit your request online by
visiting the Elsevier web site at http://elsevier.com/locate/permissions, and selecting
Obtaining permission to use Elsevier material

Notice
No responsibility is assumed by the publisher for any injury and/or damage to persons
or property as a matter of products liability, negligence or otherwise, or from any use
or operation of any methods, products, instructions or ideas contained in the material
herein. Because of rapid advances in the medical sciences, in particular, independent
verification of diagnoses and drug dosages should be made

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

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

ISBN: 978-0-444-53096-7

For information on all Elsevier publications
visit our website at books.elsevier.com

Printed and bound in The Netherlands

07 08 09 10 11 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER BOOK AID International Sabre Foundation

Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Preface

The American Institute of Chemical Engineers (AIChE) was founded in 1908. In a brief period of about one century that has passed since then, chemical engineering has gained recognition as a field of engineering. The proliferation of chemical engineering today is entirely due to the great efforts of our ancestors during these years.

Today, the scope of chemical engineering has expanded to cover an extremely wide range and it is no exaggeration to say that chemical engineering deals with almost all phenomena concerned with materials and that the subject of chemical engineering will cover an increasingly larger area in the future.

Chemical engineering has often been referred to as a study in methodology. The author, however, doubts the veracity of such a statement. Approaches in chemical engineering, as observed by the author, are determined by individual phenomena/processes, and each of these phenomena/processes is studied individually. Moreover, associated or related phenomena/processes are not considered, despite them being a part of the same chemical engineering field. In fact, the author was unable to perceive an obvious connection between such methods, and therefore, believes that chemical engineering is clearly not a study in methodology that is based on a consistent viewpoint.

The phenomena that are treated in the subject of chemical engineering can be classified into two groups:

(1) phenomena that are definite and can be expressed by formulas such as differential equations and
(2) phenomena that can be expressed only by probability terms.

The phenomena that are expressed by formulas can be explained by using basic concepts such as Newtonian mechanics; further, the scope for introducing a different way of thinking is limited. It is necessary then to focus on phenomena that are so uncertain and random that it is impossible to clarify them accurately. Even in mechanical operations in chemical engineering, there are quite a few phenomena that can be expressed by probability terms, and methods developed are specific to each operation. A good example can be found in mixing and separation operations/equipment that are representative operations/equipment in chemical engineering. Mixing and separation are phenomena that are clearly understood by the probability terms. However these phenomena are related to each other as the front and rear of an event each phenomenon has been discussed individually. For example, the evaluation indices for mixing and separation operations/equipment were defined individually, and there is no close relationship between them. In order to definitely and positively say that chemical engineering is a study in methodology, it is necessary that the above indices be defined from a common viewpoint. In other words, by considering both the phenomena from the same viewpoint, the evaluation indices described above should be discussed consistently.

In order to treat the phenomena that can be expressed only by the probability terms based on a consistent viewpoint, the phenomena should be observed through a consistent viewpoint. The author selected “information entropy” as the viewpoint. However, information entropy was not a familiar term in chemical engineering, and it was intuitively known that it has only the possibility to at least define the evaluation indices for the mixing and separation operations/equipments described above. Additionally, it was expected that other uncertain and random phenomena that are treated in chemical engineering can be discussed and understood from a consistent viewpoint by using the information entropy. If the above points are established, a new development of chemical engineering may be established, and it might become the first step for establishing a new high systemization of chemical engineering.

The author has written this book by focusing on the consistent viewpoint of “information entropy.” The author aims to present a consistent viewpoint and describe new and useful knowledge by using “information entropy.” This book is aimed at researchers and graduate students who are conducting researches in new chemical engineering fields. The author naturally recognizes that information entropy may not be the only possible viewpoint, and how the degree of information entropy is useful for the other phenomena that are described by only the terms of probability is not clarified.

The author would like to express his gratitude to a number of students who carried out researches in the author’s laboratory for their M.E. theses or B.E. theses. Finally, the author is indebted to Dr S. Ito of the Tokyo Institute of Technology for his valuable teachings. The author would also like to express his sincere appreciation for the continued moral support offered by his family: Hiroko, Mari, Chie, Aki, and Suzuka.

K. Ogawa

CHAPTER 1

Information Entropy

1.1 Introduction

The phenomena studied in chemical engineering are classified into two groups:

(1) definite phenomena that can be expressed by formulae such as differential equations,
(2) phenomena that can be expressed only by probability terms.

There is no clear scope to improve the methods of investigation of the phenomena that are expressed by formulae such as those in Newtonian mechanics. On the other hand, no two phenomena that can be expressed by probability terms are similar and as such, the methods used to investigate such phenomenon (e.g., the evaluation indices for mixing and separation operations/equipment) differ based on the nature of the phenomenon or process. In other words, there is no consistent technique for treating such phenomena that should be expressed by probability terms. The author has considered that such phenomena should be treated from a consistent viewpoint and reached to put on the glasses of information entropy to treat the phenomena. In this chapter, before discussing the main subject, the steps in the development of chemical engineering are surveyed; further, the necessity of a consistent viewpoint in chemical engineering is clarified. Next, the concept of information entropy and its important features are explained in detail. In addition, the sensitiveness of human experience for quantity is discussed in order to examine the suitability of the introduction of information entropy. It is believed that by at least comparing the expression for the amount of human feeling with that for information entropy, the suitability of the introduction of information entropy will be understood by those readers who have a strong intention to develop new fields in chemical engineering and new approaches for studying chemical engineering.

1.2 History and expectation

(1) From unit operation processes to total engineering in chemical engineering

Before agreeing to the introduction of the new way of thinking, it is necessary to understand the development process in chemical engineering. The American

Institute of Chemical Engineers (AIChE) was founded in 1908, and the Society of Chemical Engineers Japan (SCEJ) was established in 1936. In other words, almost one century had passed since chemical engineering was recognized as a part of engineering. The proliferation and widespread use of chemical engineering today is entirely due to great efforts of many ancestors during these years.

It is impossible to obtain a product by merely carrying out a chemical reaction between the materials. It is necessary to first mix the materials before chemically reacting them; further, the products must be separated after the reaction. In general, the energy expended in physical changes before and after a chemical reaction is greater than the total energy expenditure for the process. In the initial period of the chemical manufacturing industry's establishment, physical treatment depended only on experience and judgment because there was no technique available for the physical treatment. In the second half of the twentieth century, the classification and arrangement of physical treatment was attempted; further, the concept of unit operation took shape. These ideas changed the traditional way of considering a process from an individual manufacturing system to unit operation process. It can be said that chemical engineering truly began only then. Earlier, the primary aim of chemical engineering was the establishment of unit operations and their applications. Thus, in the 1940s, the rationalization for equipment design was considered, and it was clarified that unit operations were insufficient to establish the design or develop a reactor; this gave birth to reaction engineering. In the 1960s, the trend of selecting more common and basic problems in unit operations established the new chemical engineering. This movement resulted in concepts such as transport phenomena and powder technology. In the latter half of the 1960s, process engineering, process system engineering, and so on were introduced in succession in order to treat the overall process. Then, in the 1970s, the focus of chemical engineering shifted to total engineering. In chemical engineering, processes are considered as the system regardless of objects, and it has methods of expressing systems as a combination of several factors. Therefore, chemical engineering is referred to as a study in methodology.

(2) Means and aims of chemical engineering

Chemical engineering covers an extremely wide range—unit operations that have played an important role in the early stages of chemical engineering, all systems and equipment in a chemical process (from raw material feed to product supply), and biotechnology and new material engineering, which are recent subjects of interest. It is said that the warp of chemical engineering comprises the means—mass transfer, energy balance, thermodynamics, transport phenomena, reaction engineering, system engineering, and so on—and its woof comprises the aims—petrochemistry, material, energy, environment, biology, and so on. However, it is not sufficient to consider chemical engineering from the viewpoint of the warp and woof, and it is expected that hybrid structures or another viewpoint will be used in the future. The probability terms and sensitiveness of human experience for quantity are missing in the definitions of the warp and woof. Nevertheless, it is no exaggeration to say that chemical engineering deals with almost all phenomena that are concerned with materials and that the domain of chemical engineering will encompass an increasingly wider range in future.

(3) Chemical engineering as study in methodology

Chemical engineering has often been referred to as a study in methodology. The author, however, doubts the veracity of such a statement. Approaches in chemical engineering, as observed by the author, are determined by individual phenomena/processes, and each of these phenomena/processes is studied individually. Moreover, associated or related phenomena/processes are not considered, despite them being a part of the same chemical engineering field. In fact, the author was unable to perceive an obvious connection between such methods and therefore believes that chemical engineering is clearly not a study in methodology that is based on a consistent viewpoint.

The phenomena that are studied in chemical engineering can be classified into two groups:

(1) definite phenomena that can be expressed by formulae such as differential equations,
(2) phenomena that can be expressed only by probability terms.

There is no clear scope to improve the methods of investigating phenomena that are expressed by formulae such as those in Newtonian mechanics. Further, phenomena that are uncertain and random to the point that it is impossible to express them clearly are focused upon. In fact, regarding mechanical operations in chemical engineering, there are quite a few phenomena that can be expressed by probability terms; however, the methods of investigating these phenomena are specific to each operation. A good example of such operations is mixing and separation operations/equipment, which exhibit a clear identity in chemical engineering. In other words, these phenomena differentiate chemical engineering from other engineering domains. Mixing and separation phenomena are clearly expressed by probability terms. However these phenomena are considered the front and rear of an event, and each phenomenon is investigated separately. For example, the evaluation indices for mixing and separation operations/equipment are defined separately, and there is no close relationship between these indices. Therefore, in order to positively assert that chemical engineering is a study in methodology, it is necessary that the above-mentioned indices be discussed from a common viewpoint. In other words, a consistent viewpoint should be employed to define these indices such that both aspects of an event are considered.

(4) Success expectation when using information entropy viewpoint

In order to study the phenomena that can be expressed only by probability terms consistently, they should be observed from a common viewpoint. In this regard, the author has selected "information entropy" as the consistent parameter. However, although information entropy was not familiar to chemical engineering, it was intuitively known that information entropy has the possibility to define the evaluation indices for the mixing and separation operations/equipment described earlier. When the author reflected on the function of information entropy, the following expectations came to mind—the possibility of estimating the turbulent flow structure in a chemical equipment and of establishing a new scale-up rule based on the turbulent flow structure, the possibility of expressing the size distribution of particles produced by many operations, such as drops in liquid–liquid mixing, bubbles in gas–liquid mixing, crystals in crystallization, and crushed product in crushing. Additionally, it is the author's desire to express the amount of anxiety/expectation by using information entropy in order to present a method of decision-making by chemical engineers based on the amount of anxiety. If these attempts are successful, anxiety/expectation can be added as new terms in the woof of chemical engineering. In other words, it is expected that uncertain and random phenomena in chemical engineering can be expressed and explained from a consistent viewpoint by using information entropy. If the above theories are established, chemical engineering might see some new developments, and this might become the first step toward establishing a new high systemization of chemical engineering.

(5) Concrete objects to discuss from the viewpoint of information entropy

The classification into warp (means) and woof (aims) has been described earlier. Further, the *AIChE Journal* has listed the following nine major topical areas in the instruction for contributors:

(1) fluid mechanics and transport phenomena;
(2) particle technology and fluidization;
(3) separations;
(4) process systems engineering;
(5) reactors, kinetics, and catalysis;
(6) materials, interfaces, and electrochemical phenomena;
(7) thermodynamics;
(8) bio-engineering, food, and natural products;
(9) energy and environmental engineering.

Such classification into subject areas depends on the journal and it can be said that there will be as many classifications of subject areas as the number of journals. In such a case, it is difficult to establish independent areas for human experience or phenomena that can be expressed by probability terms from those given in most of the journals in chemical engineering. Human experiences and natural phenomena have been considered to be included in each area or to be not a part of chemical engineering. Although there are a number of phenomena that can be expressed by probability terms in the field of chemical engineering, we must focus on the following limited phenomena without considering the warp and woof or the above-described subject areas:

(1) Turbulent flow structure: energy spectrum function, scale up
(2) Mixing operation/equipment: mixing performance
(3) Separation operation/equipment: separation performance
(4) Micromeritics: particle size distribution

Additionally, the following term was selected to advance chemical engineering.

(5) Human experience: anxiety/expectation, decision-making

The application of information entropy to other phenomena will be easy if the applications described above are sufficiently understood.

(6) Expectation for future chemical engineers/readers

Future chemical engineers are expected to cultivate a global viewpoint of considering everything by studying extensively beyond a traditional area. Chemical engineering should create not only specialists with deep knowledge of a specific area but also engineers who can evaluate and design the systems from a global viewpoint. Global viewpoint implies the ability to harmonize not only one part but also from a wider perspective that includes humanity.

The most important facet of the book is information entropy, which was selected by the author as a viewpoint for observing the phenomena. In any case, before going to the main subject, the gist of the concept[1–4] and important matters of information entropy are explained in detail in the next section so that the rest of the contents of the book can be understood easily.

1.3 Information

"Information" is defined as "something that regardless of its form (description, rumor, etc.) and value to an individual (likes and dislikes, good and bad, merits and demerits, etc.) replaces uncertain knowledge with more certain knowledge." For example, the information that Italy won the FIFA World Cup 2006 when conveyed to an individual is done so regardless of whether the person is interested in soccer.

1.4 Amount of information

The amount of information is a technical term that expresses the magnitude of information quantitatively. The amount of information in some news is expressed by using the degree of decrease in uncertainty or increase in certainty of the knowledge provided by the news. Uncertainty of knowledge occurs because multiple results can be considered. In order to express the degree of decrease in uncertainty of knowledge provided by the news, it is natural to use the degree of decrease in the number of possible results. Therefore, the amount of information in some news when the actual result is reported can be measured by the number of possible results n before the actual result is reported:

$$I = \log n \tag{1.1}$$

This logarithmic expression is used because only it can satisfy the requirement that the amount of information should be identical whether the news is reported in entirety or in small parts. For example, as a common case, the amount of information when a die is cast is shown as follows. When a die is about to be cast, some amount of uncertainty regarding “the number on the die” will exist. When the news that “number five was obtained” is reported, the following amount of information is obtained:

$$I_5 = \log 6$$

This is because the number of faces of a die is 6 and one of them was obtained. The amount of uncertainty present before reporting the result vanishes when the news is obtained. However, the same amount of information must be obtained when the news on the outcome is reported in two parts—“odd number” and “maximum number.” The respective pieces of news are reported in this order. The amount of information in the respective pieces of news in this case is expressed as

$$I_{\text{odd}} = \log 2$$
$$I_{\text{maximum}} = \log 3$$

This is because the number is classified into two groups:

(1) odd,
(2) even.

Additionally, there are three odd numbers on a die—1, 3, and 5.

The sum of the amount of information in the case where the news is reported in two parts must be identical to the amount of information in the case where the news is reported in entirety as “number five was obtained.” This difficult requirement is perfectly satisfied by the following logarithmic expression:

$$I_{\text{odd}} + I_{\text{maximum}} = \log 2 + \log 3 = \log 6 = I_5$$

It is difficult to describe and treat the amount of information when the number of possible results becomes very large; hence, the probability that the result occurs is widely used instead of the number of results. Therefore, instead of Eq. (1.1), the following expression is widely used for the amount of information:

$$I = -\log(1/n) = -\log P \tag{1.2}$$

where P is $1/n$.

When the news that “number five was obtained” is reported, the obtainable amount of information becomes equal to

$$I_{1/6} = -\log(1/6) = \log 6$$

This is because the probability of obtaining any number from the throw of a die is 1/6. It is obvious that the amount of information is identical to that based on the number of faces of the die. When a biased die is cast, the probability of obtaining a particular number is not constant, and the individual probability of appearance should be used in Eq. (1.2).

1.5 Average amount of information before reporting the result

In the previous section, the amount of information that is obtained by reporting the result was shown. In this section, the amount of uncertainty about the result obtained before reporting the result is shown. In other words, a method of expressing the amount of uncertainty of “the result reported” is clarified. In this section, in order to make the discussion more concrete, the following uncertainties are treated:

(1) When a die is cast, “What is the number obtained?”
(2) When multiple dice are cast simultaneously, “What are the respective numbers obtained?”
(3) When a person goes to a restaurant, “Which wine does he/she order? Which side dish does he/she order?”
(4) When a person goes to a restaurant, “What is the wine that is ordered when the person orders wine ‘a’?”
(5) When a person goes to a restaurant, “What is the type of side dish that is ordered under the condition that the ordered wine will be reported?”

In the following discussion, the event/phenomenon system is divided into two categories—a single event/phenomenon system and multiple events/phenomena system. In order to place the discussion on a concrete basis, according to the previously mentioned categories, the following two cases are considered: the case of casting one die and the case of casting multiple dice at the same time.

(1) Average amount of information in the case of a single event/phenomenon system—self-entropy (When a die is cast, "What is the number obtained?")

Before the result is known, the amount of uncertainty of the result can be expressed as the average of the respective amounts of information obtained when each of the results is reported. The average amount of information is obtained as the sum of the product of the amount of information and the probability of occurrence for each result as

$$H(X) = -\sum_i P_i \log P_i \tag{1.3}$$

This average amount of information is called as "information entropy." In particular, when the system is a single event/phenomena, the information entropy is called "self-entropy." For example, the self-entropy in the case of casting an unbiased die becomes

$$\begin{aligned} H(X) &= -\frac{1}{6}\log\frac{1}{6} - \frac{1}{6}\log\frac{1}{6} - \frac{1}{6}\log\frac{1}{6} - \frac{1}{6}\log\frac{1}{6} - \frac{1}{6}\log\frac{1}{6} - \frac{1}{6}\log\frac{1}{6} \\ &= \log 6 \quad (\log_e 6 = 1.7918) \end{aligned}$$

This value is identical to that for the case where the news "'number five occurs" is given, as was described in the previous subsection. In the case of casting a biased die, which has a probability that "number one occurs" as 1/2 and the probability that other numbers occur as 1/10, the self-entropy becomes

$$\begin{aligned} H(X) &= -\frac{1}{2}\log\frac{1}{2} - \frac{1}{10}\log\frac{1}{10} - \frac{1}{10}\log\frac{1}{10} - \frac{1}{10}\log\frac{1}{10} - \frac{1}{10}\log\frac{1}{10} - \frac{1}{10}\log\frac{1}{10} \\ &= \frac{1}{2}(\log 2 + \log 10) \\ &\quad \left(\frac{1}{2}(\log_e 2 + \log_e 10) = 1.4978\right) \end{aligned}$$

The unit of information entropy depends on the base of the logarithm as follows:

$$H = -\sum_i P_i \log_2 P_i \qquad \text{[binary unit], [bit], [digit]}$$

$$H = -\sum_i P_i \log_e P_i \qquad \text{[natural unit], [nat]}$$

$$H = -\sum_i P_i \log_{10} P_i \qquad \text{[decimal unit], [dit], [Hartley]}$$

However, the unit need not be a cause of concern since the information entropy is generally discussed based on the relative values.

Let us deal with the origin of "entropy."

The word entropy was coined by Clausius in the 1850s on the basis of "energy," where "en" is the prefix and "erg," which is equivalent to "ergon," implied "work"; further, the "y" implies a "person" in Greek. In other words, energy means the person in charge of work. On the other hand, the word "entropy" comprises the prefix "en" from "energy" and "trope," which means "change" in Greek. In other words, "entropy" implies "the person in charge of the change."

"Information entropy" is named after entropy in thermodynamics since the equation of its definition, Eq. (1.3) is similar to that of entropy in thermodynamics. "Information entropy" is also termed "negentropy" because its formula has a negative sign.

(2) Average amount of information in the case of multiple events/phenomena system

When systems with multiple events/phenomena are discussed, the relationship among the events/phenomena determines the way they are considered. In order to simplify the discussion, the case of a binary system is considered.

As mentioned above, a binary system is subdivided into the following two groups:

(1) *Exclusive system*: There is no relationship among the constituent events/phenomena (e.g., the case of casting multiple dice at the same time).
(2) *Non-exclusive system*: There is some relationship among the constituent events/phenomena (e.g., the case of an order for wines and side dishes; perfectly corresponding phenomena are involved in this system.)

(a) *Combined entropy* (When multiple dice are cast simultaneously, "What is the number on the respective dice?" When a person goes to a restaurant, "What is the kind of wine that is ordered? What is the kind of side dish that is ordered?")

If the events/phenomena are mutually exclusive (e.g., the case of casting two dice simultaneously), the amount of uncertainty of the result is expressed as

$$H(X, Y) = -\sum_{i}\sum_{j} P_{ij} \log P_{ij} \tag{1.4}$$

where P_{ij} is the probability that i-result appears in the X-system and j-result appears in Y-system. This information entropy is referred to as the "combined entropy." In this case, it cannot be said that the self-entropy with respect to X-system becomes identical to Eq. (1.3) because $\sum_{j} P_{ij} = P_i$.

$$H(X) = -\sum_{i}\left\{\left(\sum_{j} P_{ij}\right)\log\left(\sum_{j} P_{ij}\right)\right\} = -\sum_{i} P_i \log P_i \tag{1.5}$$

(b) *Conditional entropy and mutual entropy* (When a person goes to a restaurant, "What is the wine that is ordered when the person ordered wine 'a'?" When a person goes to a restaurant, "What is the side dish that is ordered under the condition that the ordered wine will be reported?")

In the case that the events/phenomena in both the systems are non-exclusive (e.g., the case of ordering for wines and side dishes), the amount of uncertainty of the result in X-system under the condition that result "a" appears in Y-system is expressed by the information entropy as

$$H(X/\mathrm{a}) = -\sum_{i} P(X/\mathrm{a}) \log P(X/\mathrm{a}) \equiv -\sum_{i} P_{i/\mathrm{a}} \log P_{i/\mathrm{a}} \tag{1.6}$$

where $P_{i/\mathrm{a}} = P(X/\mathrm{a})$ is the conditional probability of appearance of the i-result in the X-system under the condition that a-result appears in the Y-system. However, a-result does not appear in the Y-system at all times, and there is a possibility that a result other than "a" appears in the Y-system. Therefore, the information entropy for the result in the X-system can be expressed as the average of the respective amounts of information entropy obtained when each result in the Y-system is reported. In other words, the amount of uncertainty on the result in the X-system under the condition that the result in the Y-system is known can be expressed by the information entropy as

$$H(X/Y) = \sum_{j} P_j H(X/j) = -\sum_{j}\sum_{i} P_j P_{i/j} \log P_{i/j} - \sum_{j}\sum_{i} P_{ij} \log P_{i/j} \tag{1.7}$$

This information entropy is referred to as "conditional entropy."

Table 1.1 Relationships among information entropies.

	Joint Entropy	Conditional Entropy	Mutual Entropy
Exclusive phenomena	$H(X, Y) = H(X) + H(Y)$	$H(X\|Y) = H(X)$ $H(Y\|X) = H(Y)$	$I(X; Y) = 0$
Non-exclusive phenomena	$H(X, Y) < H(X) + H(Y)$	$H(X\|Y) < H(X)$ $H(Y\|X) < H(Y)$	
			$I(X; Y) = H(X) - H(X\|Y)$ $= H(Y) - H(Y\|X)$
Corresponding phenomena (1:1)	$H(X, Y) = H(X) = H(Y)$	$H(X\|Y) = 0$	$I(X; Y) = H(X)$

When there is no news of the result in the Y-system, the information entropy for the result in the X-system is expressed by the self-entropy $H(X)$. In such a case, by obtaining the news "the result in Y-system will be known," the amount of information entropy decreases to that of conditional entropy $H(X/Y)$. Therefore, the amount of information obtained from the news "the result in Y-system will be known" is expressed by the balance of $H(X)$ and $H(X/Y)$ as

$$I(X; Y) = H(X) - H(X/Y) = -\sum_i P_i \log P_i + \sum_i \sum_j P_{ij} \log P_{i/j} \tag{1.8}$$

This amount of information $I(X; Y)$ is referred to as "mutual entropy."

When each event/phenomenon corresponds perfectly with another, conditional entropy has no meaning; in other words, $H(X/Y)$ takes zero value. This fact is generally understood.

A magnitude relationship among the information entropies described above is shown in Table 1.1.

1.6 Information entropy based on continuous variable

Until now, the variable in every system has been discrete; for example, the number obtained when a die is cast. As not only a discrete variable but also a continuous variable (for example, time) appears very often in chemical engineering, it is necessary to define the information entropy for continuous variable. Of course, it is possible to define the average amount of information for a system that is based on a continuous variable, for example, time. In the case based on a continuous variable t, let $p(t_i)$ and Δt be the probability density at t_i and the very small change of continuous variable, respectively. The product of $p(t_i)$ and Δt corresponds to the probability P_i, in Eq. (1.3). When this method of treatment is applied to Eq. (1.1), the equation can be developed in the following manner:

$$\begin{aligned} H(t) &= -\sum_i P_i \log P_i = -\lim_{\Delta t\to 0} \sum_i^{\infty} p(t_i)\Delta t \log\{p(t_i)\Delta t\} \\ &= -\int_0^{\infty} p(t)\log p(t)\mathrm{d}t - \lim_{\Delta t\to 0} \log \Delta t \end{aligned} \tag{1.9}$$

The second term on the right-hand side of this equation takes an infinite value regardless of the values of $p(t)$, and only the first term changes in response to the change in the probability density distribution function $p(t)$. Therefore, the information entropy based on the continuous variable is defined as

$$H(t) = -\int p(t)\log p(t)\mathrm{d}t \tag{1.10}$$

The information entropy defined by Eq. (1.10) is as important as that defined by Eq. (1.3) in chemical engineering because there are a number of phenomena that are controlled by time as a variable in other fields of engineering as well.

1.7 Probability density distribution function for the maximum information entropy

It has been said in the natural world, the aim is to achieve the maximum value of information entropy. In this section, the relationship between a probability density distribution function and the maximum value of the information entropy is discussed. In the case of a mathematical discussion, it is easier to treat information entropy $H(t)$ based on continuous variables $H(t)$ rather than the information entropy based on discrete variables $H(X)$. In the following, $H(t)$ is studied, and the probability density distribution function $p(t)$ for the maximum value of information entropy $H(t)_{\max}$ under three typical restriction conditions is shown.

(1) Range of variable t is fixed as $-R \le t \le R$

The standardized condition of the probability density distribution function is given as

$$\int_{-R}^{R} p(t)\mathrm{d}t = 1$$

In this case, the information entropy is expressed as

$$H(t) = -\int_{-R}^{R} p(t)\log p(t)\mathrm{d}t$$

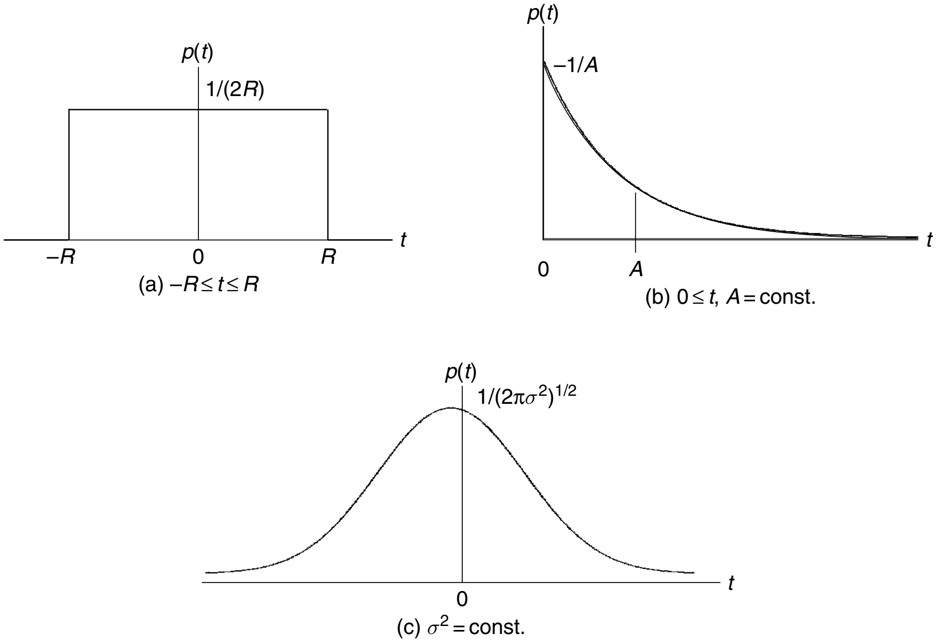
The range of integration shows the given restrictive condition. Under these conditions, the form of the probability density distribution function $p(t)$ for the maximum value of information entropy $H(t)_{\max}$ is investigated. By using calculus of variations, it is clarified that the information entropy takes the maximum value as follows:

$$H(t)_{\max} = \log(2R)$$

when the probability density distribution function $p(t)$ is given as

$$p(t) = \frac{1}{2R} \tag{1.11}$$

The probability density distribution function expressed by Eq. (1.11) is shown in Figure 1.1. This result is reasonable from the viewpoint of human experience.



(a) $-R \leq t \leq R$

(b) $0 \leq t$, $A = \text{const.}$

(c) $\sigma^2 = \text{const.}$

Figure 1.1 Probability density distribution for the maximum amount of entropy.

Let us deal with the derivation of Eq. (1.11).

The calculus of variations is applied to $H(t)$ by introducing the variable coefficient λ. When the variation method is applied to $H(t)$ as

$$\frac{\partial}{\partial p(t)} \int_{-R}^{R} \{-p(t)\log p(t)\}\mathrm{d}t + \lambda \frac{\partial}{\partial p(t)} \left\{ \int_{-R}^{R} p(t)\mathrm{d}t - 1 \right\} = 0$$

the following equation is obtained:

$$-\{1 + \log p(t)\} + \lambda = 0$$

This equation can be rewritten as

$$p(t) = \exp(\lambda - 1)$$

The variable coefficient λ should satisfy the following equation, which is obtained by substituting $p(t)$ described above into the standardized condition:

$$\exp(\lambda - 1) \cdot 2R = 1$$

Finally, it becomes clear that the probability density function $p(t)$ takes the maximum value as

$$H(t)_{\max} = \log(2R)$$

when $p(t)$ is given as

$$p(t) = \frac{1}{2R}$$

(2) Variable *t* takes a positive value, and its average value is fixed as A

The standardized conditions of the probability density distribution function and the restrictive condition are given as

$$\int_{0}^{\infty} p(t)\mathrm{d}t = 1$$

$$\int_{0}^{\infty} tp(t)\mathrm{d}t = A$$

In this case, information entropy is expressed as

$$H(t) = -\int_0^\infty p(t)\log p(t)\mathrm{d}t$$

Under these conditions, the form of the probability density distribution function $p(t)$ for the maximum value of information entropy $H(t)_{\max}$ is investigated. By using the calculus of variations, it is clarified that the information entropy $H(t)$ takes the maximum value as follows:

$$H(t)_{\max} = \log(eA)$$

when the probability density distribution function $p(t)$ is given as

$$p(t) = \frac{1}{A}\exp\left(-\frac{t}{A}\right) \tag{1.12}$$

The probability density distribution function $p(t)$ expressed by Eq. (1.12) is shown in Figure 1.1. This distribution is one of the function that is very familiar to chemical engineers.

Let us deal with the derivation of Eq. (1.12).

The calculus of variations is applied to $H(t)$ by introducing the variable coefficients λ_1 and λ_2. When the variation method is applied to $H(t)$ as

$$\frac{\partial}{\partial p(t)}\int_0^\infty\{-p(t)\log p(t)\}\mathrm{d}t + \lambda_1\frac{\partial}{\partial p(t)}\left\{\int_0^\infty p(t)\mathrm{d}t - 1\right\} + \lambda_2\frac{\partial}{\partial p(t)}\left\{\int_0^\infty tp(t)\mathrm{d}t - A\right\} = 0$$

the following equation is obtained:

$$-\{1+\log p(t)\} + \lambda_1 + \lambda_2 t = 0$$

This equation can be rewritten as

$$p(t) = \exp(\lambda_1 - 1 + \lambda_2 t)$$

The variable coefficients λ_1 and λ_2 should satisfy the following equation, which is obtained by substituting $p(t)$ described above into the first standardized condition under the assumption $\lambda_2 < 0$ as

$$-\frac{1}{\lambda_2}\exp(\lambda_1 - 1) = 1$$

λ_1 can be expressed by using λ_2 since $\exp(\lambda_1 - 1) = -\lambda_2$. Then, the probability density function $p(t)$ is expressed by using λ_2 as

$$p(t) = -\lambda_2 \exp(\lambda_2 t)$$

The variable coefficient λ_2 should satisfy the following equation, which is obtained by substituting $p(t)$ described above into the second standardized condition.

$$-\frac{1}{\lambda_2} = A$$

This result satisfies the assumption $\lambda_2 < 0$ described above. Finally, it is clear that the probability density function $p(t)$ takes the maximum value as

$$H(t)_{\max} = \log(eA)$$

when $p(t)$ is expressed as

$$p(t) = \frac{1}{A}\exp\left(-\frac{t}{A}\right)$$

(3) Variance of *t* is fixed as $\boldsymbol{\sigma}^2$

The standardized conditions of the probability density distribution function and restrictive condition are given as

$$\int_0^{\infty} p(t)\mathrm{d}t = 1$$

$$\int_{-\infty}^{\infty} t^2 p(t)\mathrm{d}t = \sigma^2$$

In this case, the information entropy is expressed as

$$H(t) = -\int_{-\infty}^{\infty} p(t)\log p(t)\mathrm{d}t$$

Under these conditions, the form of the probability density distribution function $p(t)$ for the maximum value of information entropy $H(t)_{\max}$ is investigated.

By using the calculus of variations, it is clarified that the information entropy $H(t)$ takes the maximum value as follows:

$$H(t)_{\max} = \log(2\pi e\sigma^2)$$

when the probability density distribution function $p(t)$ is given as

$$p(t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{t^2}{2\sigma^2}\right) \tag{1.13}$$

The probability density distribution function $p(t)$ expressed by Eq. (1.13) is shown in Figure 1.1. This distribution is well known as the normal distribution or Gaussian distribution.

In the chemical engineering field, there are quite a few cases in which the value of variance is discussed. For a significant discussion on the value of variance, the probability density distribution function should have the same form. Additionally, from the viewpoint of information entropy, it can be understood that the discussion based on the value of variance becomes significant when the probability density function/distribution is that for the normal distribution.

Let us deal with the derivation of Eq. (1.13).

The calculus of variations is applied to $H(t)$ by introducing the variable coefficients λ_1 and λ_2. When the variation method is applied to $H(t)$ as

$$\frac{\partial}{\partial p(t)}\int_{-\infty}^{\infty}\{-p(t)\log p(t)\}\mathrm{d}t + \lambda_1\frac{\partial}{\partial p(t)}\left\{\int_{-\infty}^{\infty}p(t)\mathrm{d}t - 1\right\}$$
$$+\lambda_2\frac{\partial}{\partial p(t)}\left\{\int_{-\infty}^{\infty}t^2 p(t)\mathrm{d}t - \sigma^2\right\} = 0$$

the following equation is derived:

$$-\{1+\log p(t)\} + \lambda_1 + \lambda_2 t^2 = 0$$

This equation can be rewritten as

$$p(t) = \exp(\lambda_1 - 1 + \lambda_2 t^2)$$

The variable coefficients λ_1 and λ_2 should satisfy the following equation that is obtained by substituting $p(t)$ described above into the first standardized condition under the assumption $\lambda_2 < 0$:

$$\left(-\frac{\lambda_2}{\pi}\right)^{1/2} = \exp(\lambda_1 - 1)$$

λ_1 can be expressed by using λ_2 since $\exp(\lambda_1 - 1) = (-\lambda_2/\pi)^{1/2}$. Then, the probability density function $p(t)$ is expressed as

$$p(t) = \left(-\frac{\lambda_2}{\pi}\right)^{1/2} \exp(\lambda_2 t^2)$$

The variable coefficient λ_2 should satisfy the following equation that is obtained by substituting $p(t)$ described above into the second standardized condition:

$$\frac{1}{2(-\lambda_2)} = \sigma^2$$

This result satisfies the assumption $\lambda_2 < 0$ described above.

Finally, it is clarified that the probability density function $p(t)$ takes the maximum value as

$$H(t)_{\max} = \log(2\pi e \sigma^2)$$

when $p(t)$ is expressed as

$$p(t) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp\left(-\frac{t^2}{2\sigma^2}\right)$$

1.8 Sensitiveness of human experience for quantity and information entropy

If the information entropy corresponds to the sensitiveness of human experience for quantity, the usefulness of the information entropy increases even more. In this section, the relationship between the information entropy and the sensitiveness of human experience for quantity is discussed. What exhibits the sensitiveness of human experience for quantity? In the daily conversations, the following expressions are widely used according to the order of the quantity size:

> "0 (zero)," "2–3 (a few)," "5–6 (several)," "10 (ten)," "20–30 (a few tens)," "50–60 (several tens)," "100 (one hundred)," "200–300 (a few hundreds),"...

The order of the quantity size described above conforms to the general usage throughout the world. This shows that the sensitiveness of human experience for quantity is incremented by one step every time an expression changes in

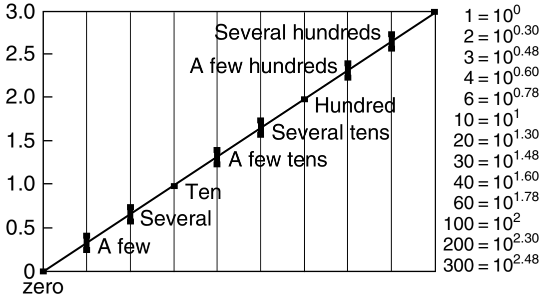


Figure 1.2 Change of sense of human experience for quantity.

the order described above; further, the difference in the sensitiveness of human experience for quantity between two neighboring steps is not very high, although the difference in the absolute value between the two steps is very large. On the contrary, it can be considered that the above-mentioned expressions were produced so that the amount of change in the sensitiveness of human experience for quantity becomes almost the same.

In this regard, the expression for the quantity is set in the same interval along the horizontal axis and each exponent of 10 of the absolute quantity that corresponds to each expression is plotted along the vertical axis, as shown in Figure 1.2. In this figure, it is definitely possible to correlate all the data by using a straight line that passes through the origin. From this figure, it is clear that the change in step is proportional to the exponent of 10 of the concrete quantity. This result means that the sensitiveness of human experience for quantity shows a logarithmic change in the response to the change in the absolute quantity.

Let us deal with the traditional knowledge concerning intensity of sensation of human and stimulus intensity.

The following two laws are well known in the field of psychology:

(1) Weber–Fechner law: $E = K\log(I/I_0)$ (E: intensity of sensation, I: stimulus, I_0: stimulus on the threshold, K: coefficient)
(2) Stevens' law: $E = C(I - I_0)^n$ or $\log E = n\log(I - I_0) + \log C$ (C: coefficient)

It can be said that the relationship between the intensity of sensation and the stimulus has nearly the same feature as that between the absolute quantity and the human experience for quantity. The relationship shown in Figure 1.2 almost corresponds to the Weber–Fechner law.

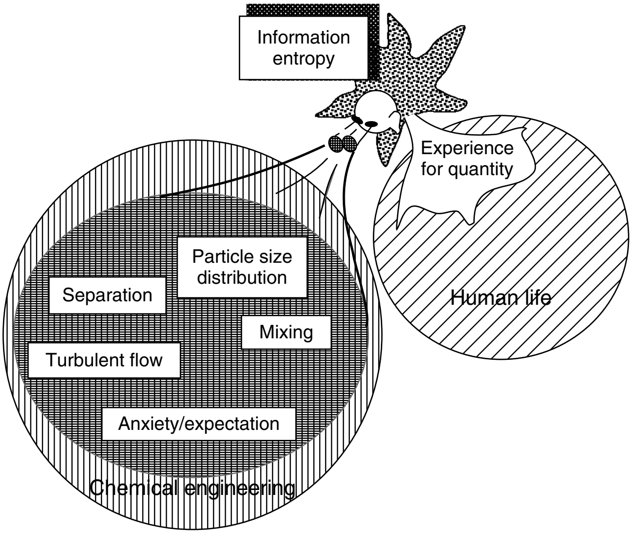


Figure 1.3 Aha! Thought.

By comparing this result with the logarithmic expression of the information entropy, it can be said that the information entropy reasonably shows a sensitiveness of human experience for quantity of information (refer to Figure 1.3).

1.9 Summary

The following issues were clarified in this chapter:

(1) In order to develop the field of chemical engineering, especially concerning the phenomena that should be expressed by the probability terms, the establishment of a consistent viewpoint is indispensable.
(2) Information entropy offers a good possibility of becoming a consistent viewpoint to treat phenomena that must be expressed by the probability terms. By using information entropy, it will become possible to define the evaluation indices for mixing and separation operations/equipment, to estimate turbulent flow structure in a chemical equipment, to establish scale-up rules based on the turbulent flow structure, to present a general particle size probability density distribution, and to define the amount of anxiety/expectation.
(3) The expression of information entropy closely approximates the expression of sensitiveness of human experience for quantity. This is a major factor for information entropy to play an important role in chemical engineering.

CHAPTER 2

Mixing Phenomena

2.1 Introduction

Since the use of the information entropy viewpoint was initially established[1] to create consistent evaluation indices for mixing and separation operations/equipment, we will first focus on mixing operations/equipment.

Mixing substances in a vessel is a technique that goes back to a time when primitive men used to cook. Subsequently, even after experiences spanning several thousand years, mixing phenomena in a stirred vessel have not been sufficiently elucidated.

Mixing phenomena in chemical equipment are accelerated by a forced flow, that is, convection and velocity fluctuation. In general, forced flow in equipment is produced by a movable part such as an impeller. Mixing phenomena are classified into two categories:

(1) macromixing,
(2) micromixing.

The scale of macromixing is one that is sufficiently discriminated by the naked eye or normal detector measurements. On the other hand, the spatial scale of micromixing corresponds to the size of a molecule. However, it is not easy to detect molecular size mixing in fluid mixing. Therefore, in chemical engineering, the mixing for sizes in the range from sufficiently larger than a molecule to clearly smaller than the equipment should be appropriately placed in the category of micromixing.

Let us deal with the features of multi-phase mixing in chemical engineering.

Liquid–liquid mixing is widely used in liquid–liquid extraction operation, dispersion polymerization, emulsion polymerization, and so on. In general, in a mutually insoluble liquid–liquid system, one liquid is dispersed into another and the liquid with the smaller volume transforms into droplets; however, when the volume of both liquids is almost identical, it is unclear which liquid becomes a droplet.

Gas–liquid mixing with aeration is a very important operation in chemical reactions (such as hydrogenation and chlorination reaction and cell culture operations).

Solid–liquid mixing is used to float solid particle in a liquid and is important for solid solutions, solid catalyst reactions, crystallization, and so on.

Solid–solid mixing is different from other multi-phase mixing with respect to the fact that the liquid phase does not participate in it. Solid–solid mixing is promoted when the outer force exceeds the gravity and the surface force. The mechanism consists of three factors:

(1) convective mixing (depending on the convective flow by rotation of equipment or impeller),
(2) shearing mixing (depending on the friction and collision that are caused by the difference in particle velocity),
(3) diffusive mixing (depending on random walk of particles that are caused by the surface state, form, size, contact state, etc., of particles).

The representative mixing equipment in the chemical industry is a stirred vessel, which has an impeller as a movable part. The stirred vessel itself is the main part of the process occasionally, and it typically operates in the background and has an inconspicuous role. However, the success or failure of the mixing operation/equipment significantly influences the success of the process. The aims of the mixing operations/equipment that utilize a group of complex mixing phenomena are classified into two groups:

(1) to homogenize,
(2) to control the rate of transport phenomena (heat/mass) or reaction.

The first aim of mixing by using impellers in liquid–liquid/liquid–gas/liquid–solid particle mixing is to disperse the droplets/bubbles/solid particles in order to increase the interface area of the two phases, continuous phase and dispersed phase, in the vessel. However, in the chemical engineering field, the most important aim is to control the rate of transport phenomena or reaction. The next section pays attention to the aim of achieving homogeneity because there has been no sufficiently consistent discussion on this aim, such as the definition of the evaluation indices of mixing operations/equipment. Further, several examples of the application of the indices to mixing operations/equipment are shown. The author expects the readers to understand the usefulness of the newly defined indices for the evaluation of mixing operations/equipment.

2.2 Index for evaluation of mixing performance

There are two standpoints on the evaluation of the mixing performance and therefore, the evaluation indices for mixing are classified into two categories according to the standpoints:

(1) mixing capacity,
(2) mixedness.

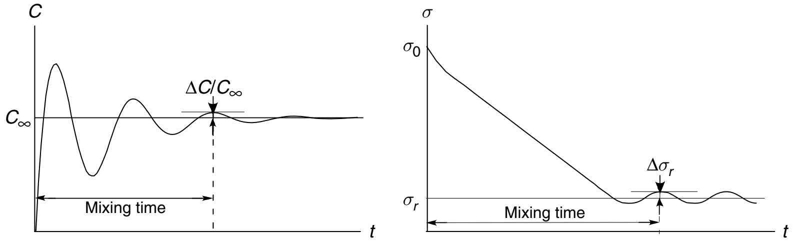
The mixing capacity is an index of the mixing ability of the equipment or operation condition, whereas the mixedness is an index of the degree of the mixing state or the spatial distribution of inner substances in the equipment. However, these indices need not be discussed separately because it is possible to discuss the mixing capacity on the basis of the change in mixedness with time or rather the mixing rate. If the aim of mixing is to homogenize, focus has often been only on the means of knowing the termination of mixing. However, if the mixing process as a function of space and time is neglected, it is difficult to discuss the most suitable equipment or operation condition for the transport phenomena or reaction. In other words, the mixing rate is an indispensable factor for discussing the most suitable equipment or operation condition for transport phenomena or reaction. As shown in Table 2.1, the traditional indices of mixing performance are classified into two groups based on the two standpoints described above.

Most of the indices of the mixing capacity in the left-hand side column in Table 2.1 are related to the mixing rate—residence time for the flow system (e.g., ratio of the standard deviation of the probability density distribution of the residence time to the average residence time; residence time is the stay time of the inner substance in an equipment), circulation time for a batch system (e.g., ratio of the standard deviation of the probability density distribution of the circulation time to the average circulation time; circulation time is the time required for one circulation of the inner substance in an equipment), mixing time (e.g., the time required for the concentration of the inner substances at a specific position in the equipment to reach a final constant value within some permissible deviation), and so on.

The index of the degree of spatial distribution of the inner substances in the equipment listed in the right-hand side column in Table 2.1 is based on the degree of difference of the state from the perfect separating state or perfect mixing state. This index has been expressed by using the ratio of the variance of the inner substance concentration distribution of the practical state to that of the perfect mixing state or perfect separating state. Obviously, standard deviation is used instead of variance later. The index is called "mixedness" or "degree of mixing" and has been used not only for liquid–liquid mixing but also for solid–solid mixing. Needless to say, it is possible to evaluate the mixing rate based on the rate of variation in mixedness with time.

Table 2.1 Traditional indices of mixing performance.

Index of Mixing Capacity	Index of Mixing State
Flow system:	
$\frac{1}{T}\left\{\int_0^\infty E(t)(t-T)^2 \mathrm{d}t\right\}^{1/2}$	$\frac{\sigma_0^2-\sigma^2}{\sigma_0^2-\sigma_r^2}$
T: average residence time	$\frac{\sigma_0-\sigma}{\sigma_0-\sigma_\mathrm{r}}$
$E(t)$: residence time density function	$\frac{\sigma_\mathrm{r}}{\sigma}$
Batch system:	
$\frac{1}{T_\mathrm{C}}\left\{\int_0^\infty E_\mathrm{C}(t)(t-T_\mathrm{C})^2 \mathrm{d}t\right\}^{1/2}$	$1-\frac{\sigma^2}{\sigma_0^2}$
T_C: average circulation time	$1-\frac{\sigma^2}{\sigma_0^2}$
$E_\mathrm{C}(t)$: circulation time density function	$\frac{\ln\sigma_0^2-\ln\sigma^2}{\ln\sigma_0^2-\ln\sigma_\mathrm{r}^2}$
	σ_0^2: standard deviation in case of complete separation state σ_r^2: standard deviation in case of final state
Mixing time:	



(a) Fluid–fluid mixing (b) Solid–solid mixing

Further, mixing/turbulent diffusivity derived by considering that the mixing process in a liquid is caused by the random movement of inner substances based on the turbulent flow can also be used as an index for the evaluation of the local mixing rate. The mass balance in the flow field is written as

$$\frac{\mathrm{D}C}{\mathrm{D}t}=\frac{\partial}{\partial x_i}\left(\varepsilon_\mathrm{C}\frac{\partial C}{\partial x_i}\right) \tag{2.1}$$

where ε_C is the mixing/turbulent diffusivity, C the concentration, and t the time. However, although this index ε_C is theoretically convenient for the development of the mixing process, it is very difficult to determine the value of ε_C practically.

In the following discussion, a new method based on information entropy is discussed to evaluate the mixing performance of operations/equipment after taking into account that the mixing process is basically a function of space and time. Mixing in the flow mixing and batch mixing systems is discussed. Further, the traditional methods to evaluate the mixing performance are also sufficiently taken into account.

Let us deal with the derivation of turbulent diffusivity.

Mass balance in a laminar flow can be written as follows:

$$\frac{\partial C}{\partial t} + U_i \frac{\partial C}{\partial x_i} = \frac{\partial}{\partial x_i}\left(D \frac{\partial C}{\partial x_i}\right)$$

Mass balance in a turbulent flow can be written as follows:

$$\frac{\partial \overline{C}}{\partial t} + \overline{U_i}\frac{\partial \overline{C}}{\partial x_i} + \frac{\partial}{\partial x_i}\overline{u_i' c'} = \frac{\partial}{\partial x_i}\left(D \frac{\partial \overline{C}}{\partial x_i}\right)$$

under the conditions of

$$C = \overline{C} + c' \quad \text{and} \quad U_i = \overline{U_i} + u_i'$$

After the terms are arranged, the following formula is obtained:

$$\frac{\partial \overline{C}}{\partial t} + \overline{U_i}\frac{\partial \overline{C}}{\partial x_i} = Ct + U_i C x_i = \frac{\partial}{\partial x_i}\left(D \frac{\partial \overline{C}}{\partial x_i} - \overline{u_i' c'}\right)$$

If cross-correlation is assumed to be in proportion to the average concentration gradient,

$$-\overline{u_i' c'} = \varepsilon_{\mathrm{C}} \frac{\partial \overline{C}}{\partial x_i}$$

Finally, the above formula becomes

$$\frac{\mathrm{D}\overline{C}}{\mathrm{D}t} = \frac{\partial}{\partial x_i}\left((D + \varepsilon_{\mathrm{C}})\frac{\partial \overline{C}}{\partial x_i}\right)$$

When it is assumed that $\varepsilon_{\mathrm{C}} \gg D$, the final formula is obtained as

$$\frac{\mathrm{D}\overline{C}}{\mathrm{D}t} = \frac{\partial}{\partial x_i}\left(\varepsilon_{\mathrm{C}}\frac{\partial \overline{C}}{\partial x_i}\right)$$

In this formula, ε_{C} is called the turbulent diffusivity.

2.3 Evaluation of mixing performance based on transition response method

There have been several attempts to use the transition response method for the evaluation of the mixing performance of operations/equipment. In the transition response method for the flow system, a tracer is injected into the inlet and the change in its concentration at the outlet with time is measured. On the other hand, in the case of the batch system, the tracer is injected into some specific position, and the change in the spatial distribution of the concentration in the equipment with time is measured. In the following discussion, a method based on information entropy to evaluate the mixing operations/equipment on the basis of the transient response method is discussed.

Let us deal with the transient response method. The transient response method is a method of measuring the characteristics of a system; in particular, this method is effective when the dynamic characteristics of the system are investigated. In order to clarify the characteristics of the system, a comparison between the input and the output wave forms is useful. In general, the following three input wave forms have been widely used (Figure 2.1):

(1) *Step response method*: The input signal is changed stepwise from the steady-state value to some specific value. This method is mainly used to discuss the dynamic characteristics of a system.

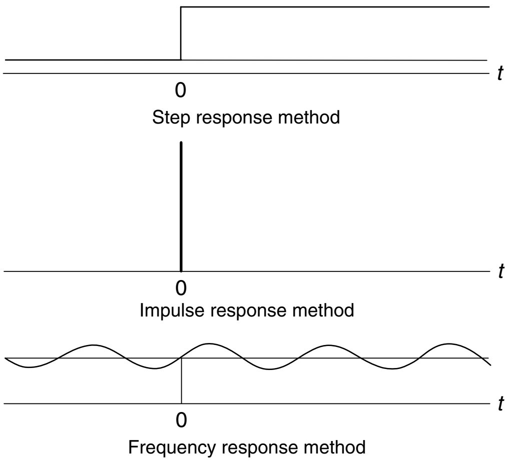


Figure 2.1 Typical transient response methods

(2) *Impulse (delta) response method*: The input signal is changed in the form of a delta function. This method is widely used in chemical engineering to investigate the residence time probability density distribution function.
(3) *Frequency response method*: The input signal is changed in the form of a trigonometrically function.

(1) Flow system

In the case of a flow system, the mixing performance of operations/equipments should be evaluated by focusing on the mixing state of inner substances at the outlet[2]. However, as a substitute for this, indices based on the standard deviation of the residence time probability density distribution (RTD) have attracted attention. In the following discussions, RTD has also been considered. The residence time is the stay time of the inner substance in equipment. Since one element requires a short time and another element requires a long time from an inlet to an outlet, the residence time does not take a constant value and it exhibits a distribution. The residence time probability function shows the fraction of the inner substance that has a residence time ranging from 0 to t. The RTD function is obtained by differentiating the residence time probability function with respect to time and arranging it so that its integral value over the whole time becomes unity. When the injection of the tracer is expressed in the form of a delta function, the concentration change in the tracer with time at the outlet becomes RTD. Of course, even if the injection of the tracer is not expressed in the form of a delta function, it is possible to obtain RTD by applying suitable mathematical steps. (Please refer to another book for the procedure.) The following discussion is based on the condition that the RTD function is already known (Figure 2.2).

When one tracer element is observed at the inlet, the mixing performance of the operation/equipment is evaluated from the viewpoint of information entropy based on the uncertainty regarding "the amount of time taken by the tracer element to reach the outlet." The time the tracer is injected is set as the origin. The definite integral from zero to infinity of the RTD function of the dimensionless residence time ($\tau(= t/T,\ T = V/Q, V$: volume of equipment, Q: flow rate)) should be unity. Additionally, by considering the definition of dimensionless time, its average value is fixed at unity. Therefore, the standardized and restrictive conditions can be written as

$$\int_0^\infty E(\tau)\mathrm{d}\tau = 1 \tag{2.2}$$

$$\int_0^\infty \tau E(\tau)\mathrm{d}\tau = 1 \tag{2.3}$$

$E(\tau)\Delta\tau$ shows the probability that the dimensionless residence time is in the range of $\tau \sim \tau + \Delta\tau$. The uncertainty regarding "the amount of time taken by

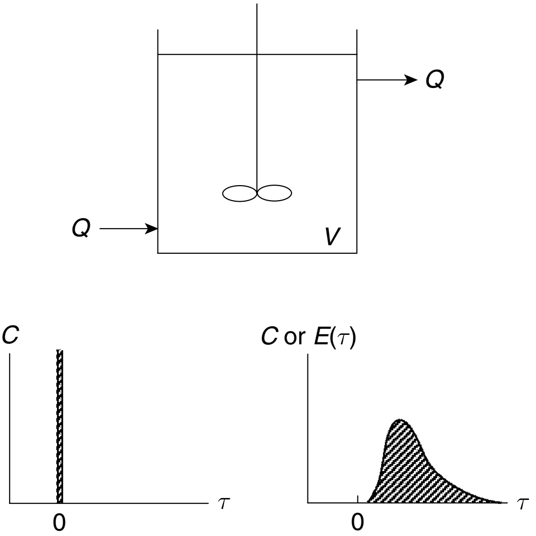


Figure 2.2 Impulse response method in a flow system.

the tracer element to reach the outlet" is expressed by the information entropy according to Eq. (1.10) as

$$H(\tau) = -\int_0^\infty E(\tau) \log\ E(\tau) \mathrm{d}\tau \tag{2.4}$$

This amount of uncertainty disappears when the residence time of the observed tracer element is known. Next, the RTD function for the maximum and minimum values of $H(\tau)$ are discussed mathematically. Since the variable τ takes a positive value and the average of τ is fixed at unity, according to Section 1.6, $H(\tau)$ assumes maximum and minimum values as

$$H(\tau)_{\max} = \log e \quad \text{at} \quad E(\tau) = \exp(-\tau) \tag{2.5a}$$

$$H(\tau)_{\min} = 0 \quad \text{at} \quad E(\tau)_{\tau \neq a} = 0 \quad \text{and} \quad E(\tau)_{\tau = a} = \infty \tag{2.5b}$$

where "a" is some specific dimensionless residence time.

The correspondence of the conditions for the maximum and minimum values of $H(\tau)$ with the practical mixing phenomena is considered. The condition under which $H(\tau)$ takes the maximum value is realized when the perfect mixing flow is established in the equipment, that is, when the concentration of tracer in the equipment is perfectly homogeneous every time and the concentration of tracer at the outlet decreases exponentially. On the other hand, the condition under which $H(\tau)$ takes the minimum value is realized when a piston flow is established in the equipment, that is, when the tracer does not disperse in the equipment completely and the concentration of the tracer at the outlet shows the form of the delta function at $\tau = 1$. From the practical viewpoint, it is impossible to consider that $E(\tau)$ becomes infinity except when $a = 1$ (that is $\tau = 1$). Therefore, the mixing capacity can be defined as the degree of approach from the piston flow to the perfect mixing flow by making use of the maximum and minimum values of $H(\tau)$ as

$$M = \frac{H(\tau) - H(\tau)_{\min}}{H(\tau)_{\max} - H(\tau)_{\min}} = \frac{-\int_0^\infty E(\tau)\log E(\tau)\mathrm{d}\tau}{\log e} \tag{2.6}$$

Let us deal with the piston flow and perfect mixing flow.

Piston flow: This flow assumes that the fluid velocity is uniform over the entire cross-section of the vessel. Each element of fluid that enters the vessel marches through single file without intermingling with other fluid elements that entered earlier or later.

Perfect mixing flow: This flow assumes that the vessel contents are completely homogeneous, and no difference exists between the various portions of the vessel, and the outlet stream properties are identical to the vessel fluid properties.

When the base of the logarithm is e, the denominator in this equation becomes unity and Eq. (2.6) becomes simpler. The mixing capacity defined by Eq. (2.6) has a value from zero for the piston flow to unity for the perfect mixing flow:

$$0 \le M \le 1 \tag{2.7}$$

Many traditional mixing indices are based on the standard deviation or the variance that shows the degree of the range of RTD. The newly defined mixing capacity above evaluates the performance of mixing on the basis of not only the range/extent of distribution but also the characteristic of the tailing parts of RTD. This point is characteristic of the newly defined mixing capacity.

It is not easy to precisely express the RTD function obtained by experiments by using a formula. Therefore, in order to calculate the mixing capacity practically, it is convenient to treat the residence time as discontinuous time, that is, as a function of the discrete time in an interval of $\Delta\tau$ (dimensionless residence time interval). From this, the probability that the observed element reaches the outlet in $\tau_i \sim \tau_i + \Delta\tau$ is shown as $E(\tau_i)\Delta\tau$, and the amount of uncertainty regarding "the amount of time taken by the observed tracer element to reach the outlet" is expressed by the information entropy as

$$\begin{aligned} H(\tau) &= -\sum_{i}^{m} E(\tau_i)\Delta\tau \log\{E(\tau_i)\Delta\tau\} \\ &= -\sum_{i}^{m} E(\tau_i)\Delta\tau \log E(\tau_i) - \log \Delta\tau \end{aligned} \tag{2.8}$$

where m is the number of discrete times that is sufficiently large to express RTD. This amount of uncertainty disappears when the residence time of the observed tracer element is known. Next, $E(\tau_i)$ for the maximum and minimum values of $H(\tau)$ are discussed mathematically. Because the second term on the right-hand side of Eq. (2.8) becomes constant, only the first term on the right-hand side end of Eq. (2.8) should be considered. This term has the same meaning as Eq. (2.4) because the time interval $\Delta\tau$ should have a very small value. In such a case, $H(\tau)$ has the following maximum and minimum values:

$$H(\tau)_{\max} = \log e - \log \Delta\tau \quad \text{at} \quad E(\tau_i) = \exp(-\tau_i) \tag{2.9a}$$

$$H(\tau)_{\min} = -\log \Delta\tau \quad \text{at} \quad E(\tau_i)_{\tau_i \neq a} = 0 \quad \text{and} \quad E(\tau_i)_{\tau_i = a} = \infty \approx 1/\Delta\tau \tag{2.9b}$$

where "a" indicates some specific dimensionless residence time. These conditions occur when the perfect mixing flow and piston flow are established identical to the case in which the residence time is considered as a continuous variable. Of course, from a practical viewpoint, it is impossible to consider that $E(\tau_i)$ becomes $1/\Delta\tau$ at a residence time other than $a = 1$. Therefore, the mixing capacity can be defined as the degree of approach from the piston flow to the perfect mixing flow by making use of the maximum and minimum values of $H(\tau)$ as follows:

$$M = \frac{-\sum_{i=1}^{m} E(\tau_i)\Delta\tau \log\{E(\tau_i)\Delta\tau\}}{\log e} \tag{2.10}$$

When the base of the logarithm is e, the denominator in the equation becomes unity, Eq. (2.10) becomes simpler. The mixing capacity defined above has a range from zero for the piston flow to unity for the perfect mixing flow:

$$0 \leq M \leq 1 \tag{2.11}$$

Let us deal with the derivation of the change in concentration with time for the perfect mixing flow from the mass balance equation (Figure 2.3).

The mass balance equation is expressed as

$$V\frac{\mathrm{d}C}{\mathrm{d}t} = -QC$$

By the definite integration of time from zero to infinity, the following relationship can be obtained:

$$C = C_0 \exp\left(-\frac{t}{V/Q}\right)$$

when the condition of $t = 0;\ C = C_0$.

When the impulse response method is used, the following standardized condition should be satisfied:

$$\int_0^\infty C\,\mathrm{d}t = 1$$

Finally, the change in concentration with time for the perfect mixing flow can be obtained as follows:

$$C = \frac{1}{V/Q}\exp\left(-\frac{t}{V/Q}\right) = \frac{1}{T}\exp\left(-\frac{t}{T}\right) \quad (T = V/Q)$$

This formula is identical to Eq. (2.5a) or Eq. (2.9a).

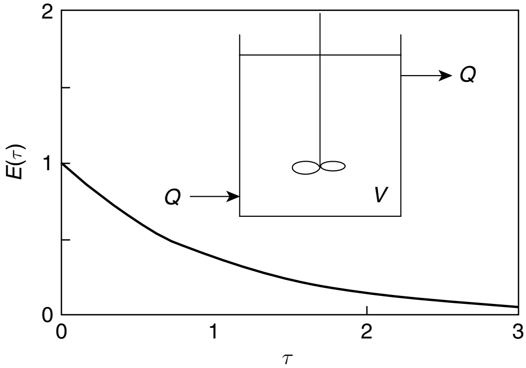


Figure 2.3 RTD curve of perfect mixing flow.

Challenge 2.1. Relationship between the mixing capacity and the number of tanks in series of perfectly mixed vessels model (a string of perfectly mixed tanks of equal size model)

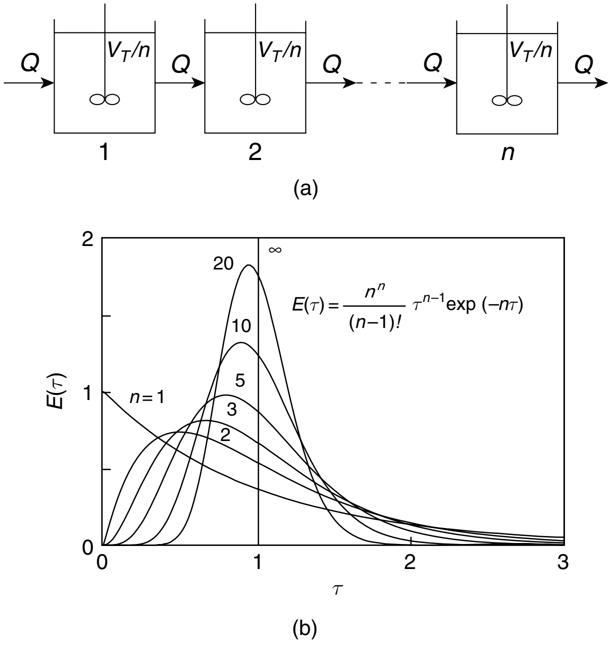
1. Scope

Since it is not easy to express the RTD function by using an accurate formula, the mixing phenomena have been frequently discussed based on mixing models that can sufficiently express RTD. One such model is the series of perfectly mixed vessels model (a string of perfectly mixed tanks of equal size model hereafter SPMV model). This model consists of a series of equivolume perfect mixing tanks as shown in Figure 2.4(a).

This model can be applied not only to a tank-type equipment but also to a tubular type equipment. Mathematically, the use of the SPMV model implies that the mixing process is expressed by a lumped parameter model and not by a distributed parameter model. The RTD function in this model is expressed as

$$E(\tau) = \frac{n^n}{(n-1)!}\tau^{n-1}\exp(-n\tau) \tag{2.12}$$

where n is the number of tanks. As the number of tanks increases, RTD changes drastically, as shown in Figure 2.4(b). By using this model, it is possible



(a)

(b)

Figure 2.4 (a) SPMV model in a flow system. (b) RTD curves in SPMV model in a flow system. (c) Mixing capacity change with number of tanks in SPMV model in a flow system.

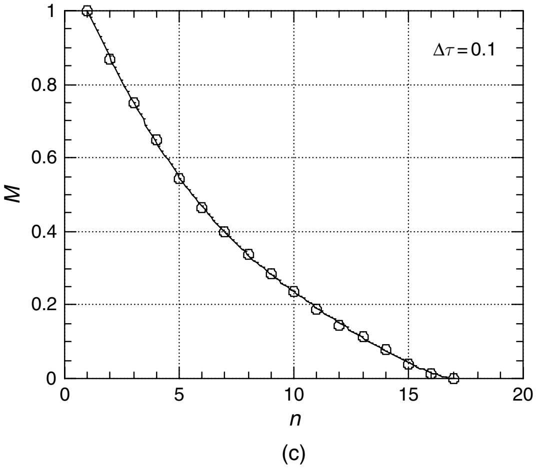


Figure 2.4 (Continued)

to express the RTD function in the flow range from perfect mixing flow to piston flow. In fact, the number of tanks for the measured RTD function in the experiments is determined by referring to this figure, and various types of analyses are possible. Therefore, clarifying the relationship between the mixing capacity and number of tanks of the SPMV model is very important.

Let us deal with the tubular-type equipment and tank-type equipment.

The tubular-type equipment and tank-type equipment are defined as follows:

(1) *Tubular type equipment*: Long and slender straight pipe or a coiled pipe or a U-shaped curved pipe. The flow direction is considered to be only along the axial direction.
(2) *Tank type equipment*: Typical vessel. The flow direction is considered to be three dimensional.

In addition, there are a lumped and distributed parameter models.

The premise is that the system characteristic can be expressed by using either of the parameter in the following manner:

(1) *Lumped parameter model*: The parameter is concentrated at a finite point.
(2) *Distributed parameter model*: The parameter is continuously distributed in the system.

2. Aim

To clarify the relationship between the mixing capacity M defined by Eq. (2.10) and the number of tanks n of the SPMV model.

3. Calculation

(a) Condition

(i) Model: SPMV model:

$$E(\tau) = \frac{n^n}{(n-1)!}\tau^{n-1}\exp(-n\tau)$$

(ii) The dimensionless residence time interval: $\Delta\tau = 0.1$.

(b) Method
By changing the number of tanks n, the mixing capacity M is calculated based on Eq. (2.10).

4. Calculated result

Figure 2.4(c) (Mixing capacity M versus number of tanks n).

5. Noteworthy point

(a) The mixing capacity M becomes less than 0.9 when the number of tanks n becomes greater than 2.
(b) On the other hand, when the number of tanks n is close to 15, the flow in the equipment is understood to be the piston flow.

6. Supplementary point

(a) There are many models in addition to the SPMV model described above. In all the models, it is possible to calculate the mixing capacity if RTD is expressed by a formula.

Challenge 2.2. Relationship between mixing capacity and impeller rotational speed and set of positions of the inlet and outlet in a flow-stirred vessel

1. Scope

Although flow-stirred vessels have been used widely, the issue of “the best position for the inlet and outlet in order to establish effective mixing” is still unsolved. Additionally, the effect of impeller rotational speed on the mixing capacity has not been clarified. Therefore, it is necessary to clarify the best position for the inlet and outlet and the effect of the impeller rotational speed on the mixing capacity.

2. Aim

To clarify the relationships between the mixing capacity M defined by Eq. (2.10) and impeller rotational speed N and to investigate the effect of the positions of the inlet and outlet on the mixing capacity M defined by Eq. (2.10).

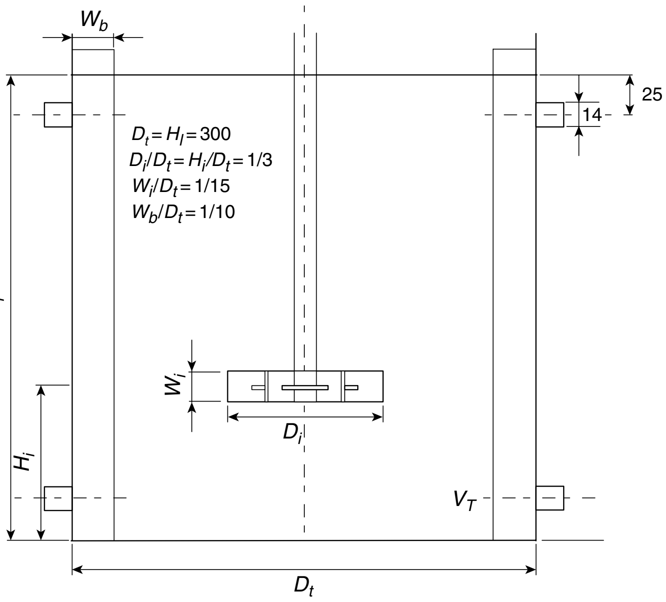
3. Experiments

(a) Apparatus

Stirred vessel: Figure 2.5(a) (Cylindrical flat bottom vessel, four baffles).
Impeller: six-flat blade turbine type.

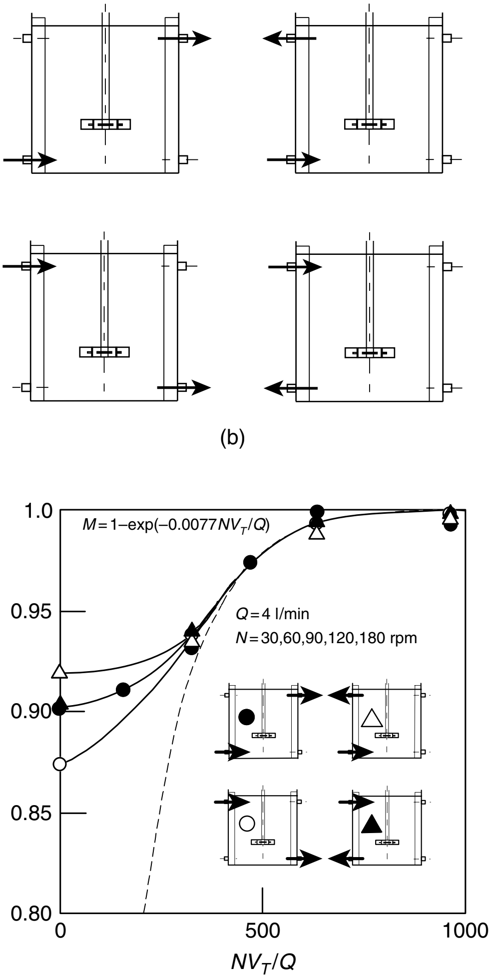
(b) Condition

Fluid: ion-exchanged water
Tracer: KCl saturated solution (2 ml).



(a)

Figure 2.5 (a) Stirred vessel of a flow system. (b) Four sets of positions of inlet and outlet of a flow system. (c) Mixing capacity change with impeller rotational speed for four sets of positions of inlet and outlet of a flow system.



(b)

(c)

Figure 2.5 (Continued)

Set of positions of the inlet and outlet: Figure 2.5(b) (four sets of positions of the inlet and outlet. Both the inlet and outlet are arranged in the same vertical cross-section through the impeller axis in every set).

Flow rate (Q): 3.5 and 4 l/min.

Impeller rotational speed (N): 30, 60, 90, 120, and 180 rpm ($Re = 0.5 \times 10^4$–3×10^4).

(c) Tracer measurement method

Electrode conductivity probe.

(d) Procedure

The tracer is injected at the inlet by using an injector in the form of an impulse, and the concentration of the tracer at the outlet is measured by the electrode conductivity probe. Based on the change in concentration with time, the mixing capacity M defined by Eq. (2.10) is calculated.

4. Experimental results

Figure 2.5(c) (Mixing capacity versus dimensionless time for each set of positions of the inlet and outlet when $Q = 4 \mathrm{l/min}$).

5. Noteworthy point

(a) M is close to 0.9 only by allowing flow regardless of the positions of the inlet and outlet. The highest value of M is in the case where the inlet position is set at the lower part of the wall and the outlet is set at a higher part of the same side wall.
(b) In the range of $N(V_T/Q) > 400$, regardless of the inlet and outlet positions, M increases gradually with N according to some specific exponential function.

6. Supplementary point

(a) Under other experimental conditions, the tendency of the results described above remains almost the same.

(2) Batch system

The evaluation of the mixing capacity of a batch mixing system has been widely performed on the basis of the mixing time. The mixing time is defined as the time required for the concentration of the tracer at some specific position in the equipment to reach a final constant value within some permissible deviation. For instance, since every fluid element is definitely considered to pass through the impeller position, this impeller position is typically selected as some specific position. The change in the concentration with time of the tracer that is injected into the impeller position is measured at the same position. Besides, the circulation time probability density distribution (CPD) function is sometimes utilized to evaluate the mixing capacity. The circulation time is defined as the time required for a fluid element at some specific position in the equipment to circulate through the equipment. On the other hand, the mixing state, or rather the degree of dispersion of the substances in a vessel, has been evaluated by the indices based on the variance of the spatial concentration distribution of the tracer or on the scale of the tracer concentration inhomogeneity of the tracer or the change in those indices with time.

In this section, the mixing performance of a batch system is discussed according to the following steps. First, a tracer element is observed. Second, the mixing

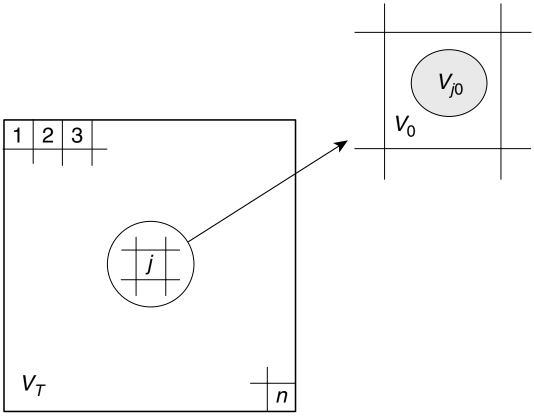


Figure 2.6 Definition diagram for batch system-I.

state is evaluated from the viewpoint of information entropy based on the uncertainty regarding "the region in which the observed tracer element is at time t." The time when the tracer is injected is set as the origin. Third, the method of evaluation of the mixing capacity based on the change in the mixing state with time is discussed. In order to define the mixedness that shows the degree of the mixing state quantitatively, the following conditions are set (Figure 2.6):

(1) The vessel (volume V_T) is partitioned into n imaginary regions with equal volume V_0, which is equal to the total volume of the tracer:

$$nV_0 = V_T \tag{2.13}$$

(2) The volume of the tracer in the j-region at time t is V_{j0}:

$$\sum_j V_{j0} = V_0 \tag{2.14}$$

(When the concentration of the tracer is used in place of the volume of the tracer, the following equation can be used:

$$\sum_j C_{j0} V_0 = C_0 V_0 \tag{2.15}$$

where C_0 and C_{j0} are the concentration of the tracer before the beginning of mixing and that in j-region at time t, respectively.)

Under such conditions, the uncertainty regarding "the region in which the observed tracer element is at time t" is considered. Since the volume of the tracer in the j-region at time t is V_{j0}, the probability that the tracer is in the j-region at time t is V_{j0}/V_0. Therefore, the amount of information that is obtained by the news that the tracer element observed is present in the j-region is expressed as

$$I(R_j) = -\log \frac{V_{j0}}{V_0}$$

It is very important to note that that V_0 implies not only the volume of each region but also the total volume of the tracer. The probability that the news described above is conveyed is V_{j0}/V_0; in other words, the probability is equal to the ratio of the tracer volume in the j-region to the total volume of the tracer. Therefore, the amount of uncertainty regarding "the region in which the observed tracer element is at time t" is expressed by information entropy, which is the average amount of the information, described above, of all possible news:

$$H(R) = \sum_j^n \frac{V_{j0}}{V_0} I(R_j) = -\sum_j^n \frac{V_{j0}}{V_0} \log \frac{V_{j0}}{V_0} \equiv \sum_j^n P_{j0} \log P_{j0} \tag{2.16}$$

This amount of uncertainty disappears when the region in which the observed tracer element is present is known. Next, the probability function P_{j0} for the maximum and minimum values of $H(R)$ is discussed mathematically. Since the range of the variable j is fixed as $1 \leq j \leq n$, according to Section 1.7, $H(R)$ takes the maximum and minimum values as follows:

$$H(R)_{\max} = -\log \frac{V_0}{V_T} = \log n \quad \text{at} \quad P_{j0} = \frac{V_0}{V_T} = \frac{1}{n} \tag{2.17a}$$

$$H(R)_{\min} = 0 \qquad \text{at} \quad P_{j0,j\neq a} = 0 \quad \text{and} \quad P_{j0,j=a} = 1 \tag{2.17b}$$

where "a" is some specific region.

The correspondence of the conditions for the maximum and minimum values of $H(R)$ with the practical mixing phenomena is considered. The condition under which $H(R)$ takes the maximum value is realized when the perfect mixing state is established, that is, when the tracer disperses in the equipment homogeneously and the concentration of the tracer at every region becomes identical. On the other hand, the condition under which $H(R)$ takes the minimum value is realized when a state of no mixing is maintained, that is, when the tracer does not disperse in the equipment at all or moves to another region without breakup. Therefore, mixedness can be defined as the degree of approach from the no-mixing state to the perfect-mixing state by using the maximum and minimum values of $H(R)$ as follows:

$$M = \frac{H(R) - H(R)_{\min}}{H(R)_{\max} - H(R)_{\min}} = \frac{-\sum_j^n P_{j0} \log P_{j0}}{\log n} \tag{2.18}$$

The mixedness defined by Eq. (2.18) has values from zero for the no-mixing state to unity for the perfect-mixing state:

$$0 \le M \le 1 \tag{2.19}$$

In practice, it is not easy to measure the concentration of the tracer in all the extremely small regions at the same time. Then, in general, the equipment is partitioned into suitable volume regions in which the mixing can be assumed to be sufficiently perfect. However, in this case, each partitioned region differs from other regions with respect to the volume; by considering that each partitioned region is a collection of small equivolume V_0 regions, as described above, the same treatment that was described above is possible. In other words, there is no difference between the result for the case where the equipment is partitioned into suitable volume regions and the case where the equipment is partitioned into equivolume V_0 regions. However, note that the method of summation must be available.

As observed, it becomes possible to quantitatively express the degree of the mixing state based on the spatial distribution of the concentration of tracer in the equipment. Accordingly, the evaluation of the mixing capacity of the operation/equipment becomes possible by using the change in mixedness with time.

Challenge 2.3. Change in mixedness with time of three typical impellers[3]

1. Scope

There are various purposes of using a stirred vessel, and the required mixing effect depends on the purpose. For any mixing purpose, rapid and homogeneous dispersion is required. In a stirred vessel, forced convection by the rotation of an impeller occurs, that is, each element of the fluid has an individual velocity; finally, turbulent flow based on the shear stress accelerates the mixing. Therefore, the shape of the impeller has a very important effect on the mixing state. However, there are various types of impeller shapes, and traditional impellers are classified into three types:

(1) disc turbine type
(2) turbine type
(3) pitched type.

Recently, in contrast to the traditional small impellers, a large impeller has been manufactured in Japan. In general, it has been said that the disc turbine-type, turbine-type, and pitched-type impellers mainly produce radial, tangential, and axial direction flows, respectively. However, the type of impeller that is most effective for rapid and homogeneous dispersion has not yet been clarified. Therefore, this is an issue that requires immediate attention.

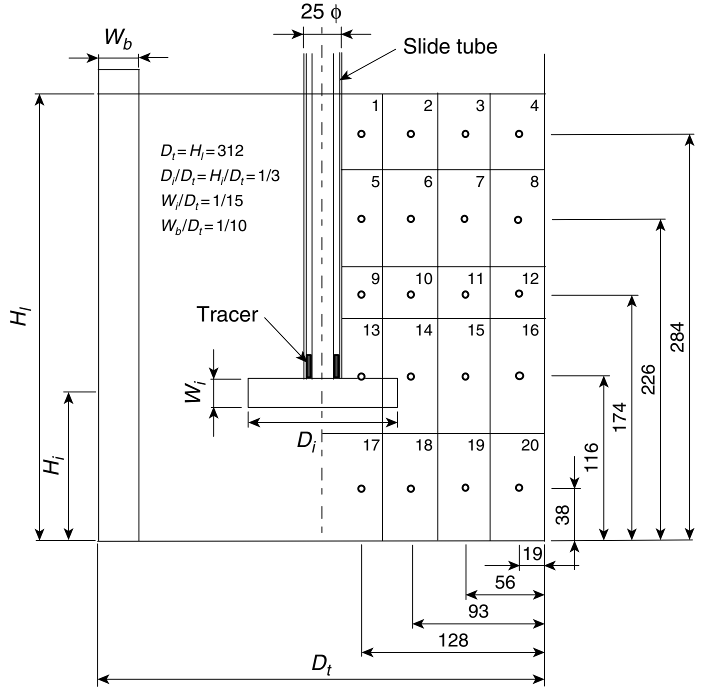
2. Aim

To clarify the relationship between the mixedness *M* defined by Eq. (2.18) and time *t* for the three typical impellers—disc turbine type, turbine type, and pitched type impellers.

3. Experiments

(a) Apparatus

Stirred vessel: Figure 2.7(a) (Cylindrical flat bottom vessel, four baffles).
Impeller: Figure 2.7(b)



(a)

Figure 2.7 (a) Stirred vessel of a batch system and imaginary partition of vessel. (b) Three types of impeller. (c) Relationship between mixedness and real time of FBDT impeller in a stirred vessel. (d) Relationship between mixedness and dimensionless time of FBDT in a stirred vessel. (e) Relationship between mixedness and dimensionless time of FBT and 45° PBT in a stirred vessel.

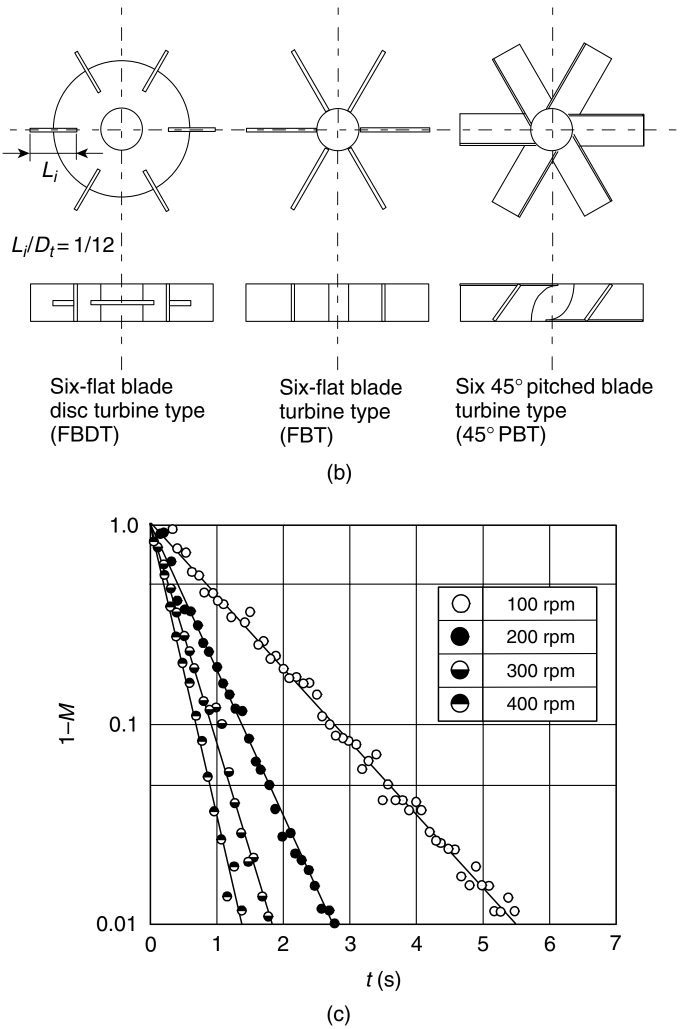
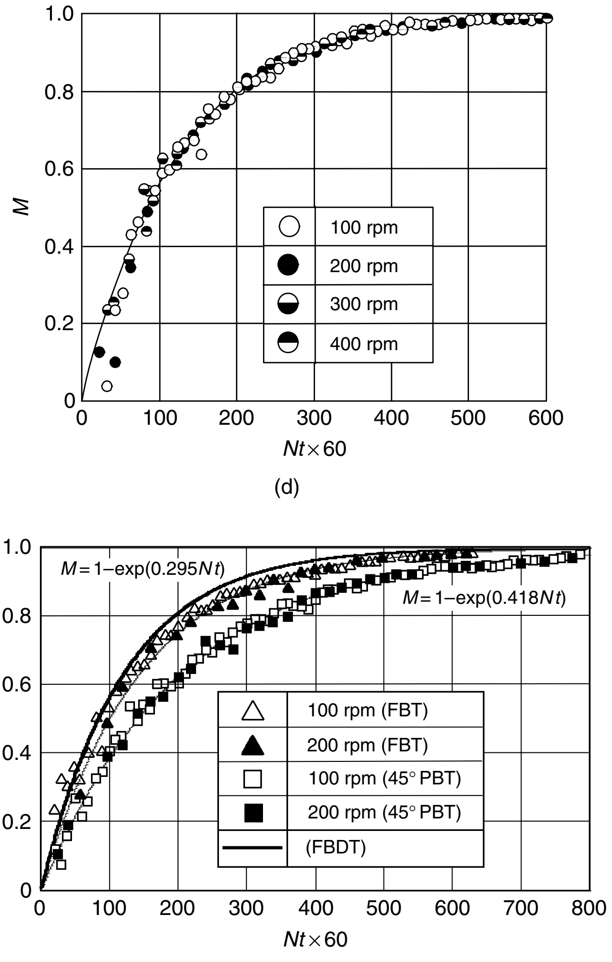


Figure 2.7 (Continued)

six-flat blade disc turbine-type impeller (FBDT impeller)
six-flat blade turbine-type impeller (FBT impeller)
six 45° pitched blade turbine-type impeller (decent flow; 45° PBT impeller)

(b) Condition

Fluid: ion-exchanged water
Tracer: 0.1 mole/l KCl solution (2 ml).
Impeller rotational speed (*N*): 100, 200, 300 and 400 rpm ($Re = 1.80 \times 10^4$–7.21×10^4)
Partition of stirred vessel: Figure 2.7(a)



(d)

(e)

Figure 2.7 (Continued)

(The imaginary partition of the vessel is performed by considering the flow pattern for each impeller as each region can be considered to be a complete mixing state regardless of whether the impeller styles and number of regions are suitable for measuring the concentration of the tracer.)

(c) Tracer measurement method

Electrode conductivity probe.

(d) Procedure

The electrolyte aqueous solution, as the tracer, is placed in the clearance between the impeller axis and the outer cylinder. After confirming that the flow in a vessel has a steady state under a fixed impeller rotational speed, the outer cylinder is instantaneously moved up to inject the tracer into the vessel. From the time of tracer injection, the concentration of the tracer is measured in every partitioned imaginary region by using an electrode conductivity probe, and the mixedness M defined by Eq. (2.18) is calculated. (When a transient response method is used to discuss the mixing performance, the best position for injecting the tracer is the impeller position; this is because in a stirred vessel, every element of the fluid passes through the impeller position at least. Therefore, calculating the mixedness based on the spatial distribution of concentration of tracer in the vessel by using Eq. (2.18) is suitable by injecting the tracer into the impeller position[3]).

4. Experimental results

Figure 2.7(c) (Mixedness versus time for FBDT impeller).
Figure 2.7(d) (Mixedness versus dimensionless time for FBDT impeller).
Figure 2.7(e) (Mixedness versus dimensionless time for FBT impeller and 45° PBT impeller).

5. Noteworthy point

(a) When the $1-M$ is plotted versus t on a log-normal graph paper, a straight line for each impeller rotational speed is obtained. This tendency is identical to that of the other impeller styles.

(b) When M is plotted versus $Nt \times 60$ (N: impeller rotational speed [rpm], t: time [s]), all the data of M are on a unique line regardless of the impeller rotational speeds. This tendency is the same as that of the other impeller styles.

The relationship between M and $Nt \times 60$ for the three impeller styles are written as follows:

$$\text{FBDT impeller}: \quad M = 1 - \exp(-0.498Nt) \tag{2.20a}$$

$$\text{FBT impeller}: \quad M = 1 - \exp(-0.418Nt) \tag{2.20b}$$

$$45^{o}\,\text{PBT impeller}: \quad M = 1 - \exp(-0.295Nt) \tag{2.20c}$$

(c) It is possible to state that the basic process of mixing is the same regardless of the impeller styles; this is because the change in M with t can be expressed by the same formula regardless of the impeller styles. In other words, the fact that the common variable is Nt regardless of the impeller styles shows that the product of the average circulation time T_c(= (tank volume)/(discharge flow rate)= $V_T/(\pi D^3 N)$) and the impeller rotational speed N determines the mixing state. (The formula can be derived based on the model that the change in the mixing rate with time is proportional to the difference of M from unity, or in other words, the difference of the mixing state from the final state.)

(d) When each M in the above equations is differentiated with respect to t in order to compare the mixing rate of each impeller, the following equations are obtained:

$$\text{FBDT impeller}: \qquad \frac{\mathrm{d}M}{\mathrm{d}t} = 0.498N\,\exp(-0.498Nt) \tag{2.21a}$$

$$\text{FBT impeller}: \qquad \frac{\mathrm{d}M}{\mathrm{d}t} = 0.418N\,\exp(-0.418Nt) \tag{2.21b}$$

$$45^\circ\,\text{PBT impeller}: \qquad \frac{\mathrm{d}M}{\mathrm{d}t} = 0.295N\,\exp(-0.295Nt) \tag{2.21c}$$

When these three equations are compared, the order of the mixing rate under a constant impeller rotational speed becomes

$$\text{FBDT impeller} > \text{FBT impeller} > 45^\circ\ \text{PBT impeller}$$

The order of the mixing capacity of each impeller style is considered to be identical to that of the mixing rate described above. Therefore, the order of the mixing capacity of each impeller style is considered to be the same as that of the mixing rate.

$$\text{FBDT impeller} > \text{FBT impeller} > 45^\circ\ \text{PBT impeller}$$

(e) For achieving the value of $Nt \times 60$ for reaching any amount of M, the 45° PBT impeller requires a larger t than the FBDT impeller. (For example, for M to reach 0.9, the value for the FBDT impeller becomes $Nt \times 60 = 300$. However, for the 45° PBT impeller, the value becomes $Nt \times 60 = 450$, that is, 1.5 times the value for the FBDT impeller.)

6. Supplementary point

(a) It is expected that a different result can be obtained by partitioning the vessel into smaller regions, thereby increasing the number of regions. However, in these experiments, increasing the number of regions did not appear to cause a considerable difference in the results.

(b) It is confirmed[4] that the same formula with regard to change in M with t could be obtained when the tracer was injected into the center of the liquid surface in the 3.1 m I.D. vessel ($V_T = 280\,\mathrm{m}^3$) in which the three stages of the FBDT impeller were set[4].

(c) It is very significant that the result described above is obtained when the tracer is injected into a certain impeller position, and the same result may not be obtained if the tracer is injected into another position.

Challenge 2.4. Relationship between mixedness and impeller rotational speed in an aerated stirred vessel

1. Scope

The aim of the operation of an aerated stirred vessel is classified into two groups:

(1) to obtain homogeneous and stable dispersion by dispersing fine bubbles in the liquid phase,
(2) to promote mass transfer or reaction between gas and liquid.

The aerated stirred vessel is used as a gas–liquid mixing equipment, when high gas hold-up and gas–liquid mass transfer rate are required. In the operation of the aerated stirred vessels, it has been found that the flooding phenomenon occurs when the power input by an impeller is too low to disperse a certain gas flow rate. Below the flooding point, the gas dispersion is inefficient so that both the gas hold-up and gas–liquid mass transfer rate decrease sharply. This is not a desirable condition and must be avoided in the gas–liquid mixing operation. In order to satisfy this requirement, it is extremely important to sufficiently disperse fine bubbles in the entire vessel. The influence of the impeller rotational speed on the mass transfer rate can be divided into two regions—region without agitation effect and region with agitation effect.

Let us deal with regions with and without agitation effect (refer to Challenge 5.3).

A region without agitation effect and that with agitation effect are defined as follows:

(1) *Region without agitation effect*: At a very low impeller rotational speed, up to a certain minimum impeller rotational speed, the transfer rate does not improve with stirring. In this region, the mass transfer rate depends on the gas load and type of sparger but is independent of the impeller rotation speed.
(2) *Region with agitation effect*: As the impeller rotational speed exceeds the minimum speed, the mass transfer rate increases quickly and linearly with the increasing impeller rotational speed.

The mixing performance of the stirred vessel was discussed based on the change in concentration of the injected gas with time at some specific position in the vessel. However, although the aerated stirred vessel has been widely used, its mixing performance has still not been discussed in detail. Therefore, clarifying the relationship between the mixedness and impeller rotational speed in an aerated stirred vessel is important.

2. Aim

To clarify the relationship between mixedness M defined by Eq. (2.18) and impeller rotational speed N in an aerated stirred vessel.

3. Experiments

(a) Apparatus

Stirred vessel: Figure 2.8(a) (Flat cylindrical bottom vessel, four baffles).
Impeller: FBDT impeller.

(b) Condition

Fluid:

Liquid phase: ion exchange water.
Gas phase: nitrogen gas.

Tracer gas: carbonic acid gas.
Impeller rotation speed (N): 200, 300, 400 rpm ($Re = 1.07 \times 10^5$–2.14×10^5).
Gas flow rate: $5.17 \times 10^{-5}\ \mathrm{m^3/s}$.
Partition of stirred vessel: similar to Figure 2.7(a).

(c) Typical measurement method

Electrode conductivity probe.

(d) Procedure

The impeller is rotated at a fixed speed, and nitrogen gas is fed at a fixed flow rate. After it has been confirmed that the flow in the vessel has reached a steady state and that the carbonic acid gas in the water has been desorbed, carbonic acid gas is fed as the tracer for 2 s at the same flow rate as the fixed flow rate of the nitrogen gas by changing the valve. After injection of the tracer gas, nitrogen gas is fed at the fixed flow rate again by changing the valve. The change in the concentration of tracer with time in each region is measured by making use of the electrode conductivity probe, and the mixedness M defined by Eq. (2.18) is calculated.

4. Experimental results

Figure 2.8(b) (Mixedness versus dimensionless time).

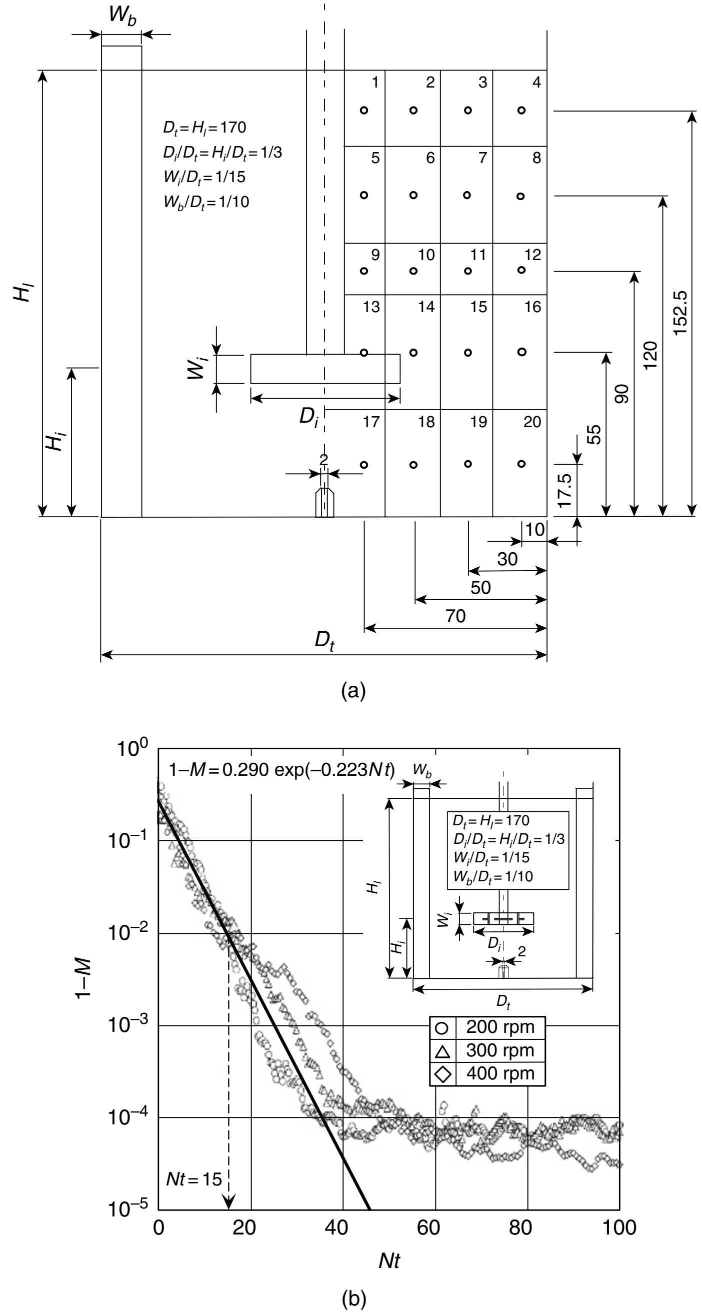


Figure 2.8 **(a) Aerated stirred vessel and imaginary partition of vessel. (b) Relationship between mixedness and dimensionless time in an aerated stirred vessel.**

5. Noteworthy point

(a) At $Nt \sim 0$, M is biased at an almost constant value regardless of the value of N. The reason for this is estimated to be as follows. The time required for the tracer to move from the nozzle to the impeller is very short, and there is very little difference in the change in M with N in this time interval.
(b) In the region of $50 < Nt$, $1 - M$ is about $1 - M < 10^{-4}$ regardless of the value of N. This implies that the mixing is effected by the fluid motion that is completed at about $Nt \sim 50$.
(c) In the region of $0 < Nt < 15$, the change in M with t is almost the same regardless of the value of N, and the relation can be expressed as

$$1 - M = 0.290 \exp(-0.223Nt)$$

This equation means that the change in M with t depends on N; further, the difference in the flow state in the vessel that is controlled by the discharge flow rate from the impeller affects the change in mixedness with time.

6. Supplementary point

(a) However, at $Nt \sim 15$, $1 - M$ reaches a low value of about 10^{-2}, and it is confirmed that small concentration patches still remain in the vessel.

Challenge 2.5. Relationship between change in mixedness with distance along axial direction and tracer injection position along radial direction in a circular pipe[5]

1. Scope

A pipe is considered an arrangement to transfer fluids from one equipment to another and not a proper equipment by itself. Recently, much progress has been achieved in mixing or reaction in a pipe line, along with material. However, the position where the second substance should be fed for the establishment of expected mixedness over the shortest axial distance has not yet been clarified. Therefore, it is very important to clarify the spatial distribution of the concentration of tracer when it is injected into an arbitrary radial position. This is because if the spatial distribution of the concentration of the tracer is obtained, it becomes possible to calculate the mixedness at an arbitrary distance in the axial direction; in other words, the change in mixedness with distance along the axial direction can be obtained.

2. Aim

To clarify the relationship between the mixedness M defined by Eq. (2.18) and the distance in the axial direction z when the tracer is injected into the center of the pipe, and to compare it with that of the case when the tracer is injected into the wall ring region of the pipe.

3. Calculation

It is now required to obtain the spatial distribution of the concentration of the tracer for two cases—center injection and wall ring injection. In general, it is difficult to theoretically obtain the spatial distribution of the concentration of tracer, and the number of experimental results with regard to this subject is insufficient because a very long pipe is required for performing experiments. However, it is possible to estimate the distribution of the tracer concentration based on the mean velocity profile, the intensity of velocity fluctuations, and so on, during the turbulent flow in a circular pipe.

(a) Condition

Flow state: fully developed turbulent flow.

Turbulent statistical characteristics: the mean velocity profile, the intensity of velocity fluctuations, and so on (the result at $Re = 3.5 \times 10^4$–10^5 by Lawn[7] and the result at $Re = 3.8 \times 10^5$ by Laufer[8]).

Mixing phenomena: two-dimensional phenomena (axial and radial directions) in which the substance in the partitioned region disperses immediately and the concentration of the substance becomes homogeneous in the region.

Tracer injection position: center of the pipe and wall ring of the pipe. (The injection speed of the tracer is set to the average velocity as the injection region.)

Partition of the circular pipe: horizontal cross-section of the pipe is divided into 10 imaginary concentric doughnut-type regions (width is 1/10th the radius of the circular pipe); its length in the axial direction is the distance a fluid element flows with a mean velocity in 1 s.

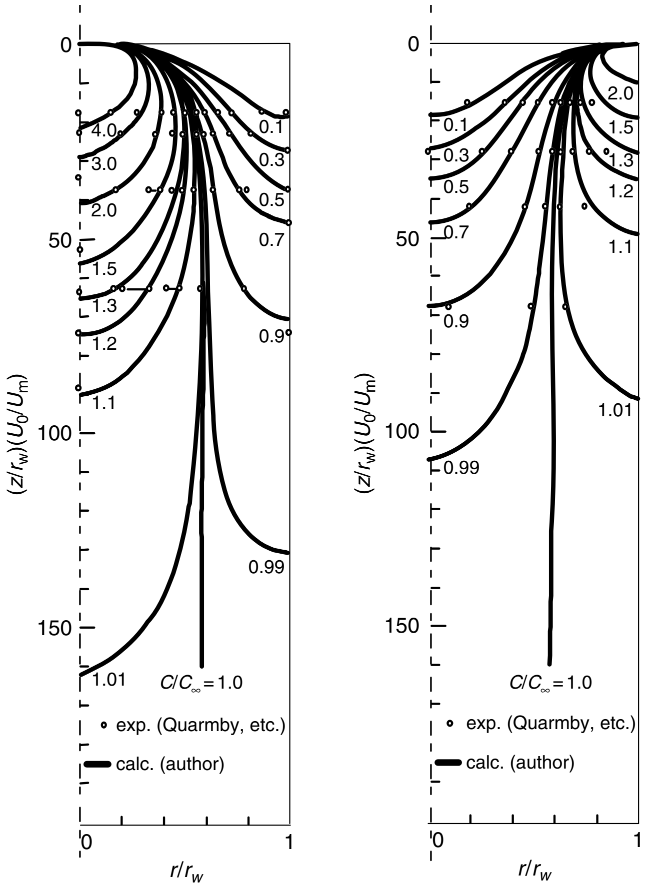
(b) Method

First, the distribution of the concentration of the tracer is estimated based on the turbulent statistical characteristics such as the mean velocity profile and the intensity of velocity fluctuations. (The method for obtaining the distribution of the tracer concentration is given in detail in the original paper.[5] Second, mixedness M based on the distribution of the concentration of the tracer in the cross-section at an arbitrary distance along the axial direction is calculated by using Eq. (2.18).

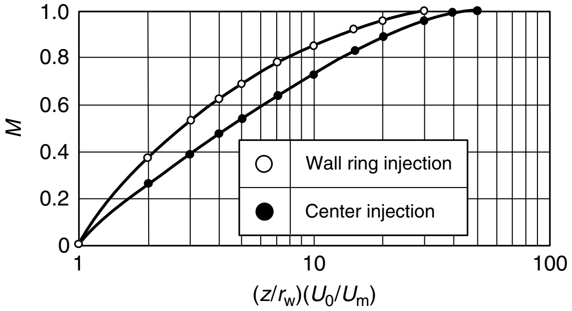
4. Calculated result

Figure 2.9(a) and (b) (Concentration distribution of the tracer in the cross-section through the axis in a circular pipe[6]).

Figure 2.9(c) (Mixedness versus distance in the axial direction).



(a) Center injection (b) Wall ring injection



(c)

Figure 2.9 (a, b) Concentration distribution of tracer in a cross-section through axis in a circular pipe (left-hand side dotted lines in both figures are pipe axis, left-hand side figure is center injection, right-hand side figure is wall ring injection; *z*: axial position, *r*: radial position, r_w: pipe radius, U_0: pipe center average velocity, U_m: cross-average velocity). (c) Mixedness change in axial direction in a circular pipe.

5. Noteworthy point

(a) It is possible to sufficiently estimate the distribution of the tracer concentration based on the mean velocity, the intensity of velocity fluctuations, and so on, because the distribution of the tracer concentration coincides sufficiently with the experimental results[6].
(b) It is clarified that for the distance required to attain any amount of mixedness, the wall ring injection requires a shorter distance than the center injection. (For example, the center injection requires about 13 times the pipe inner diameter for reaching $M = 0.9$, and it is about 1.44 times the distance of that for the case of the wall ring injection.)
(c) The mixing rate becomes as follows:
pipe wall ring injection > pipe center injection

6. Supplementary point

(a) It is confirmed that if the cross-area is divided into smaller regions, the results are not considerably different from those for the case described above.

Challenge 2.6. Relationship between mixedness and gas flow rate in a bubble column

1. Scope

The bubble column reactors represent contactors in which a gas or a mixture of gases is distributed in the liquid at the column bottom by an appropriate distributor and moves upwards in the form of bubbles causing intense mixing of the liquid phase. The aim of a bubble column is to control the rate of mass transfer and reaction between a gas and a liquid. In general, both liquid and gas are fed continuously in a countercurrent or parallel flow style. A large quantity of the gas is unsuitable for the bubble column because the pressure drop due to the gas flow becomes fairly large. However, absorption controlled by the liquid phase is suitable for a bubble column. The mixing performance of the bubble column has been discussed based on the change in concentration of the injected tracer gas with time at some specific position in the column. Although the bubble column has been widely used, a detailed discussion on its mixing performance remains incomplete. Therefore, it is significant to clarify the relationship among the mixedness, time, and gas flow rate in a bubble column.

2. Aim

To clarify the relationship among the mixedness M defined by Eq. (2.18), time t, and specific gas velocity u_g in a bubble column.

3. Experiments

(a) Apparatus

Bubble column: Figure 2.10(a) (cylindrical flat bottom vessel).

Gas distributor: (thickness 5 mm, holes 2 mmϕ × 31, regular triangle setting).

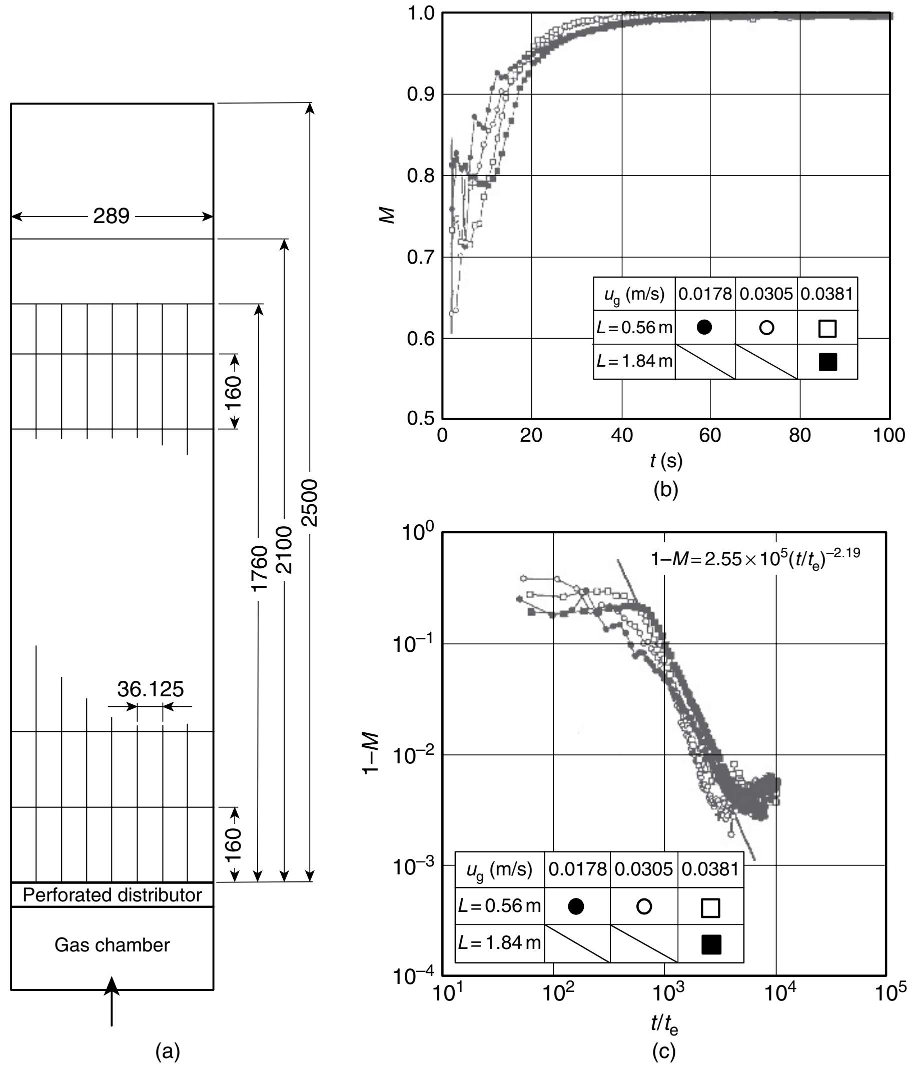


Figure 2.10 (a) Bubble column and imaginary partition of column. (b) Relationship between mixedness and real time in a bubble column. (c) Relationship between mixedness and dimensionless time based on the contact time of bubble and liquid in a bubble column.

(b) Condition

Fluid: ion exchange water.
Gas: nitrogen.
Tracer gas: carbonic acid gas.
Gas flow rate: $5.17 \times 10^{-5}\,\mathrm{m}^3/\mathrm{s}$.
Partition of bubble column: Figure 2.10(a) (Concentric doughnut-type regions).

(c) Tracer measurement method

Electrode conductivity probe.

(d) Procedure

Nitrogen gas is fed at a fixed flow rate. After confirming that the flow in the column has a steady state and that the carbonic acid gas in the water is desorbed, the carbonic acid gas as the tracer is fed stepwise at the same flow rate as the fixed flow rate of nitrogen gas by changing the valve. The change in concentration of the tracer with time in each region is measured by making use of the electrode conductivity probe after the tracer gas is fed, and the mixedness M defined by Eq. (2.18) is calculated.

4. Experimental results

Figure 2.10(b) (Mixedness versus real time).
Figure 2.10(c) (1–M versus dimensionless time t/t_e (t_e: contact time of bubble and liquid = average bubble diameter/ascend velocity of bubble = d_b/u_g)).

5. Noteworthy point

(a) In the region $t < 50\,\mathrm{s}$, M is clearly affected by u_g and the liquid height L, and mixing rate $\mathrm{d}M/\mathrm{d}t$ attains a higher value directly proportional to the small value of u_g.
(b) However, for a small liquid depth, $\mathrm{d}M/\mathrm{d}t$ reaches a higher value until $t \sim 20\,\mathrm{s}$; this tendency is reversed at $t > 20\,\mathrm{s}$.
(c) In the region $t > 60\,\mathrm{s}$, the effect of the operation condition does not appear in the result. This implies that t_e is useful in investigating the mixing phenomena in a bubble column.
(d) The relationship between M and t/t_e can be sufficiently expressed by the following formula regardless of the values of u_g and L.

$$1 - M = 2.55 \times 10^5 \left(\frac{t}{t_e}\right)^{-2.19}$$

2.4 Evaluation of mixing performance based on transition probability of inner substance

The evaluation of the mixing performance described in the previous section is only for the case where the tracer is injected into some specific position. Since the conclusion is not based on all the results that are obtained by changing the tracer injection position one by one, the scope of the application is limited. In order to establish a more universal evaluation, the movement of an inner substance from one region to another in the vessel should be considered[9].

The role of the movement of a substance among the regions in a vessel can be classified into two groups according to the direction of movement:

(1) the role of the distributor based on the outflow,
(2) the role of the blender based on the inflow.

In order to evaluate the mixing capacity, the two movements should be treated individually. In the following discussion, some element of the substance in the j-region is observed, and the mixing capacity is evaluated from the viewpoint of information entropy based on the uncertainty regarding "the region that the observed element flows in unit time" or "the region that the observed element flows in from in unit time." In order to define the mixing capacity based on the movement of an inner substance, the following conditions are set (Figure 2.11):

(1) The vessel (volume V_T) is partitioned into n imaginary regions with equal volume V_0:

$$nV_0 = V_T \tag{2.22}$$

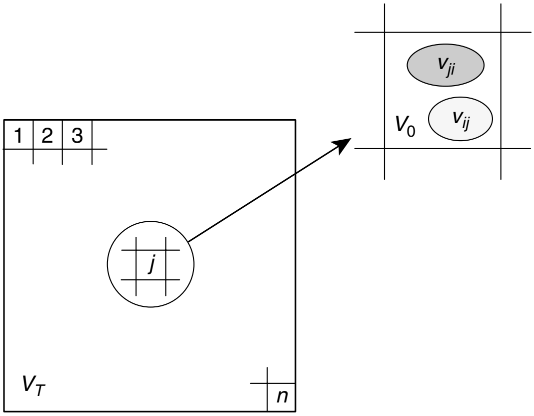


Figure 2.11 Definition diagram for batch system-II.

(2) The movement of the substance in the *j*-region is classified into two groups:

  (i) the movement as a distributor (from the *j*-region to the *i*-region),
  (ii) the movement as a blender (from the *i*-region to the *j*-region).

There is no difference in the role of mixing between the two movements. The volume of the substance that flows out from the *j*-region to the *i*-region in unit time is set to v_{ij}, while the volume of the substance that flows in from the *i*-region to the *j*-region in unit time is set as v_{ji}. It is assumed that the phenomena result in a perfect event system, that is, the substances in all the positions of the *j*-region/*i*-region are distributed to all the positions of the *i*-region/*j*-region with equal probability:

$$\sum_j v_{ij} = V_0, \quad \sum_i v_{iji} = V_0 \tag{2.23}$$

Under these conditions, uncertainty regarding "the region that the observed element flows out to in unit time" or regarding "the region that the observed element flows in from in unit time" is considered. Since the ratio of the volume that flows out from the *j*-region to the volume of the *j*-region is v_{ij}/V_0 and the ratio of the volume that flows in from the *i*-region to the volume of the *j*-region is v_{ji}/V_0, these ratios have the probability that the observed element flows out to and in from the *i*-region, respectively. Therefore, the amount of information that is obtained by the news that informs that the observed element flows out or in from the *i*-region is expressed as

$$I_{\mathrm{O}j}(R_j) = -\log\frac{V_{ij}}{V_0}$$

$$I_{\mathrm{I}j}(R_j) = -\log\frac{V_{ji}}{V_0}$$

The probabilities that the news described above are reported are V_{ij}/V_0 and V_{ji}/V_0, respectively. Therefore, the amount of uncertainty regarding "the region that the observed element flows out to in unit time" or "the region that the observed element flows in from in unit time" is, respectively, expressed by the information entropy as

$$H_{\mathrm{O}j}(R) = \sum_i^n \frac{v_{ij}}{V_0} I_{\mathrm{O}j}(R_i) = -\sum_i^n \frac{v_{ij}}{V_0}\log\frac{v_{ij}}{V_0} \equiv -\sum_i^n P_{ij}\log P_{ij} \tag{2.24a}$$

$$H_{\mathrm{I}j}(R) = \sum_i^n \frac{v_{ji}}{V_0} I_{\mathrm{I}j}(R_i) = -\sum_i^n \frac{v_{ji}}{V_0}\log\frac{v_{ji}}{V_0} \equiv -\sum_i^n P_{ji}\log P_{ji} \tag{2.24b}$$

These amounts of uncertainties disappear when the region to which the observed element flows out or the region from which the observed element flows in is known. The average of the two information entropies described above—$H_{\mathrm{O}j}(R)$ and $H_{\mathrm{I}j}(R)$—is written as

$$H_{\mathrm{L}j}(R) = \frac{H_{\mathrm{O}j}(R) + H_{\mathrm{I}j}(R)}{2} = \frac{-\sum_{i}^{n}\{P_{ij}\log P_{ij} + P_{ji}\log P_{ji}\}}{2} \tag{2.25}$$

Next, the probability function P_{ij} or P_{ji} for the maximum and minimum values of $H_{\mathrm{O}j}(R)$ and $H_{\mathrm{I}j}(R)$ is discussed mathematically. Since the range of the variable i is fixed as $1 \leq i \leq n$, according to Section 1.8, $H_{\mathrm{O}j}(R)$ and $H_{\mathrm{I}j}(R)$ take the maximum and minimum values as

$$\begin{aligned} &H_{\mathrm{O}j}(R)_{\max} = H_{\mathrm{I}j}(R)_{\max} = H_{\mathrm{L}j}(R)_{\max} = -\log(V_0/V_T) = \log n \\ &\quad \text{at} \quad P_{ij} = P_{ji} = V_0/V_T = 1/n \end{aligned} \tag{2.26a}$$

$$\begin{aligned} &H_{\mathrm{O}j}(R)_{\min} = H_{\mathrm{I}j}(R)_{\min} = H_{\mathrm{L}j}(R)_{\min} = 0 \\ &\quad \text{at} \quad P_{ij,i\neq a} = P_{ji,j\neq b} = 0 \quad \text{and} \quad P_{ij,i=a} = P_{ji,j=b} = 1 \end{aligned} \tag{2.26b}$$

where “*a*” and “*b*” are some specific regions.

The correspondence of the conditions for the maximum and minimum values of $H_{\mathrm{O}j}(R)$, $H_{\mathrm{I}j}(R)$, and $H_{\mathrm{L}j}(R)$ with the practical mixing phenomena is considered. The condition for $H_{\mathrm{O}j}(R)$, $H_{\mathrm{I}j}(R)$, and $H_{\mathrm{L}j}(R)$ to take the maximum value is realized when the substance in every region flows out equally to all regions and flows in equally from all regions. In other words, this condition occurs when a perfect mixing flow is established. On the other hand, the condition for $H_{\mathrm{O}j}(R)$, $H_{\mathrm{I}j}(R)$, and $H_{\mathrm{L}j}(R)$ to take the minimum value is realized when the fluids in every region do not disperse to any other region or move to another region without breakup. In other words, this condition occurs when piston flow is established. Therefore, the local mixing capacity for the j-region can be defined as the degree of approach from piston flow to perfect mixing flow by making use of the maximum and minimum values of $H_{\mathrm{O}j}(R)$ or $H_{\mathrm{I}j}(R)$ or $H_{\mathrm{L}j}(R)$ as follows:

(1) Index as distributor:

$$M_{\mathrm{O}j} = \frac{H_{\mathrm{O}j}(R) - H_{\mathrm{O}j}(R)_{\min}}{H_{\mathrm{O}j}(R)_{\max} - H_{\mathrm{O}j}(R)_{\min}} = \frac{-\sum P_{ij}\ \log\ P_{ij}}{\log\ n} \tag{2.27a}$$

(2) Index as blender:

$$M_{\mathrm{I}j} = \frac{H_{\mathrm{I}j}(R) - H_{\mathrm{I}j}(R)_{\min}}{H_{\mathrm{I}j}(R)_{\max} - H_{\mathrm{I}j}(R)_{\min}} = \frac{-\sum P_{ji} \log P_{ji}}{\log n} \tag{2.27b}$$

(3) Average of indices:

$$M_{\mathrm{L}j} = \frac{H_{\mathrm{L}j}(R) - H_{\mathrm{L}j}(R)_{\min}}{H_{\mathrm{L}j}(R)_{\max} - H_{\mathrm{L}j}(R)_{\min}} = \frac{-\frac{1}{2}\left(\sum P_{ij} \log P_{ij} + \sum P_{ji} \log P_{ji}\right)}{\log n} \tag{2.27c}$$

Each local mixing capacity defined by Eq. (2.27) has a value from zero for the no-mixing state to unity for the perfect-mixing state:

$$0 \le M_{\mathrm{O}j} \le 1 \tag{2.28a}$$

$$0 \le M_{\mathrm{I}j} \le 1 \tag{2.28b}$$

$$0 \le M_{\mathrm{L}j} \le 1 \tag{2.28c}$$

The information entropy for the whole vessel, $H_{\mathrm{OW}}(R)$ or $H_{\mathrm{IW}}(R)$ or $H_{\mathrm{LW}}(R)$, is obtained by taking the average of the information entropy ($H_{\mathrm{O}j}(R)$ or $H_{\mathrm{I}j}(R)$ or $H_{\mathrm{L}j}(R)$) of all regions as

$$H_{\mathrm{OW}}(R) = \sum_{j}^{n} \frac{V_0}{V_T} H_{\mathrm{O}j}(R) = -\frac{1}{n} \sum_{j}^{n} \sum_{i}^{n} P_{ij} \log P_{ij} \tag{2.29a}$$

$$H_{\mathrm{IW}}(R) = \sum_{j}^{n} \frac{V_0}{V_T} H_{\mathrm{I}j}(R) = -\frac{1}{n} \sum_{j}^{n} \sum_{i}^{n} P_{ji} \log P_{ji} \tag{2.29b}$$

$$H_{\mathrm{LW}}(R) = \sum_{j}^{n} \frac{V_0}{V_T} H_{\mathrm{L}j}(R) = -\frac{1}{2n} \sum_{j}^{n} \sum_{i}^{n} (P_{ij} \log P_{ij} + P_{ji} \log P_{ji}) \tag{2.29c}$$

The amount of information from the news regarding the outward flow of the observed element from the j-region to the i-region is the same as the amount of information from the news regarding the inward flow of the observed element from the i-region to the j-region when attention is focused on the i-region. In other words, $H_{\mathrm{O}j}(R)$, $H_{\mathrm{I}j}(R)$, and $H_{\mathrm{L}j}(R)$ take the same value: $H_{\mathrm{OW}}(R) = H_{\mathrm{IW}}(R) = H_{\mathrm{LW}}(R) \equiv H_{\mathrm{W}}(R)$. The maximum and minimum values of these information entropies are discussed mathematically. The maximum values as well as the minimum values and the conditions required for achieving them are the same as those for $H_{\mathrm{O}j}(R)$ or $H_{\mathrm{I}j}(R)$ or $H_{\mathrm{L}j}(R)$ case. Therefore, the whole mixing capacity index can be defined as the degree of approach from piston flow to perfect mixing flow by utilizing the maximum and minimum values of the information entropy, $H_{\mathrm{W}}(R)$, as follows:

$$\begin{aligned} M_{\mathrm{W}} = \frac{H_{\mathrm{W}}(R) - H_{\mathrm{W}}(R)_{\min}}{H_{\mathrm{W}}(R)_{\max} - H_{\mathrm{W}}(R)_{\min}} &= \frac{-(1/n)\sum_{j}^{n}\sum_{i}^{n} P_{ij} \log P_{ij}}{\log n} \\ &= \frac{-(1/n)\sum_{j}^{n}\sum_{i}^{n} P_{ji} \log P_{ji}}{\log n} \\ &= \frac{-\{1/(2n)\}\sum_{j}^{n}\sum_{i}^{n} (P_{ij} \log P_{ij} + P_{ji} \log P_{ji})}{\log n} \end{aligned} \tag{2.30}$$

The whole mixing capacity defined above takes a value from zero for piston flow to unity for perfect mixing flow:

$$0 \le M_W \le 1 \tag{2.31}$$

Needless to say, the whole mixing capacity can also be derived by taking the average of the local mixing capacity ($M_{\mathrm{O}j}(R)$ or $M_{\mathrm{I}j}(R)$ or $M_{\mathrm{L}j}(R)$) in all the regions.

In practice, it is not easy to obtain the movement volume of a substance at each small region in the equipment. Then, in general, the equipment is partitioned into suitable volume regions in which the mixing can be assumed to be perfect. In this case, by considering that each partitioned region is the collection of the small equivolume regions described above, the same treatment that is described above is possible. In other words, the case where the equipment is divided into suitable volume regions and the case where the equipment is divided into small equivolume regions do not show a difference in results, although the method of summation should be available.

The mixing capacity for the case where only the inflow from a specific o-region to each region occurs is a modified index that is obtained by setting $I = 0$ and then multiplied by V_T/V_0 in Eq. (2.30). The expression of the modified index becomes identical to the definition of the mixedness in the case of the transient response method in the previous section. This is only the contact point with the mixedness based on the transient response method. The CPD can be calculated if the transient probability of the inner substance is known.

Until now, a quantitative evaluation of the local mixing capacity and whole mixing capacity based on the movement of the inner substances among regions has been shown.

Challenge 2.7. Local and whole mixing capacities of FBDT impeller and 45° PBT impeller

1. Scope

In the previous section, the difference in the mixing rate among the impeller styles was clarified as

$$\text{FBDT impeller} > \text{FBT impeller} > 45^\circ\ \text{PBT impeller}$$

when the tracer is injected into the impeller position. However, there is no guarantee that the same result will be obtained when the tracer is injected into another region. Therefore, the question regarding "the impeller that has the higher value of the mixing capacity" must be answered.

2. Aim

To clarify the relationship between the mixing capacities $M_{\mathrm{O}j}(R), M_{\mathrm{I}j}(R)$, $M_{\mathrm{L}j}(R)$ and M_{w} defined by Eqs. (2.27) and (2.30), respectively, and the impeller styles.

3. Calculation

(a) Condition

Stirred vessel: Figure. 2.7(a) (Cylindrical flat bottom vessel, four baffles).
Impeller: Figure 2.7(b) (FBDT impeller and 45° PBT impeller).
Fluid: ion exchange water.
Impeller rotational speed (N): 200 rpm ($Re = 3.61 \times 10^4$).
Time interval: 0.2 s.
Partition of stirred vessel: Figure 2.7(a).
Unit volume (V_0): 100 cm^3.

(b) Method

The transition probabilities P_{ij} and P_{ji} in 0.2 s for both the impellers are estimated based on the experimental data in Challenge 2.3 (the change in the spatial distribution of the tracer concentration with time when the tracer is injected into an impeller position under an impeller rotational speed of 200 rpm) by using a computer. Based on the transition probabilities, the local mixing capacity and whole mixing capacity of both the impellers are calculated by using Eqs. (2.27) and (2.30) under the condition of $V_0 = 100\,\mathrm{cm}^3$.

4. Calculated result

Table 2.2 (Transition probabilities of two impellers).
Figure 2.12 (Contour line maps of the local mixing capacities on the semi-vertical cross-section of the vessel for each impeller style).

Table 2.2 Transition probability from j-region to i-region of FBDT impeller and 45° PBT impeller.

FBDT $(j \rightarrow i)$

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$	$i=16$	$i=17$	$i=18$	$i=19$	$i=20$
$j=1$	0.587	0	0	0	0.413	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$j=2$	0.159	0.780	0	0	0.062	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$j=3$	0	0.102	0.801	0	0	0.076	0.019	0.002	0	0	0	0	0	0	0	0	0	0	0	0
$j=4$	0	0	0.036	0.916	0	0	0.045	0.003	0	0	0	0	0	0	0	0	0	0	0	0
$j=5$	0	0	0	0	0.427	0	0	0	0.112	0.461	0	0	0	0	0	0	0	0	0	0
$j=6$	0	0	0	0	0	0.808	0	0	0.070	0.116	0.006	0	0	0	0	0	0	0	0	0
$j=7$	0	0.039	0	0.004	0	0.046	0.759	0	0	0.017	0.135	0	0	0	0	0	0	0	0	0
$j=8$	0	0	0.110	0.081	0	0	.041	0.769	0	0	0	0	0	0	0	0	0	0	0	0
$j=9$	0	0	0	0	0	0	0	0	0.686	0	0	0	0.249	0.065	0	0	0	0	0	0
$j=10$	0	0	0	0	0	0	0	0	0	0.449	0.001	0	0.020	0.530	0	0	0	0	0	0
$j=11$	0	0	0	0	0	0	0	0	0	0.121	0.756	0	0	0.106	0.017	0	0	0	0	0
$j=12$	0	0	0	0	0	0	0.095	0.277	0	0	0.048	0.580	0	0	0	0	0	0	0	0
$j=13$	0	0	0	0	0	0	0	0	0	0	0	0	0.711	0.112	0.177	0	0	0	0	0
$j=14$	0	0	0	0	0	0	0	0	0	0	0	0	0	0.378	0.439	0.182	0	0	0	0
$j=15$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.459	0.541	0	0	0	0
$j=16$	0	0	0	0	0	0	0	0	0	0	0.012	0.232	0	0	0	0.423	0	0	0	0.334
$j=17$	0	0	0	0	0	0	0	0	0	0	0	0	0.079	0.370	0	0	0.551	0	0	0
$j=18$	0	0	0	0	0	0	0	0	0	0	0	0	0.015	0	0.285	0	0.180	0.518	0	0
$j=19$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.033	0.140	0	0.307	0.520	0
$j=20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.351	0.649

(*Continued*)

Table 2.2 (Continued)

45° PBT ($j \rightarrow i$)

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$	$i=10$	$i=11$	$i=12$	$i=13$	$i=14$	$i=15$	$i=16$	$i=17$	$i=18$	$i=19$	$i=20$
$j=1$	0.459	0.541	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$j=2$	0.095	0.603	0.066	0	0.235	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$j=3$	0	0.043	0.435	0.390	0	0	0	0.132	0	0	0	0	0	0	0	0	0	0	0	0
$j=4$	0	0	0.237	0.445	0	0	0	0.318	0	0	0	0	0	0	0	0	0	0	0	0
$j=5$	0.266	0.030	0	0	0.217	0	0	0	0.086	0.401	0	0	0	0	0	0	0	0	0	0
$j=6$	0.011	0	0.141	0	0.055	0.437	0.311	0	0.013	0.034	0	0	0	0	0	0	0	0	0	0
$j=7$	0	0.070	0.108	0.175	0	0.254	0.181	0	0	0.213	0	0	0	0	0	0	0	0	0	0
$j=8$	0	0	0	0.143	0	0	0.291	0.566	0	0	0	0	0	0	0	0	0	0	0	0
$j=9$	0	0	0	0	0	0	0	0	0.252	0	0	0	0.748	0	0	0	0	0	0	0
$j=10$	0	0	0	0	0.013	0.149	0.208	0	0.095	0.263	0	0	0	0.272	0	0	0	0	0	0
$j=11$	0	0	0	0	0	0.023	0.035	0.025	0	0.028	0.328	0.033	0	0.175	0.353	0	0	0	0	0
$j=12$	0	0	0	0	0	0	0.067	0.003	0	0	0.282	0.421	0	0	0.226	0	0	0	0	0
$j=13$	0	0	0	0	0	0	0	0	0.211	0	0	0	0.516	0	0	0	0.044	0.229	0	0
$j=14$	0	0	0	0	0	0	0	0	0.008	0.063	0	0	0.014	0.509	0	0	0	0.176	0.229	0
$j=15$	0	0	0	0	0	0	0	0	0	0	0.128	0.377	0	0.028	0.278	0	0	0.070	0.120	0
$j=16$	0	0	0	0	0	0	0	0	0	0	0.046	0.029	0	0	0.209	0.614	0	0	0	0.102
$j=17$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.685	0.315	0	0
$j=18$	0	0	0	0	0	0	0	0	0	0	0	0	0	0.046	0	0	0.109	0.466	0.378	0
$j=19$	0	0	0	0	0	0	0	0	0	0	0	0	0	0.050	0.028	0	0	0.005	0.422	0.494
$j=20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.021	0.407	0	0	0.041	0.531

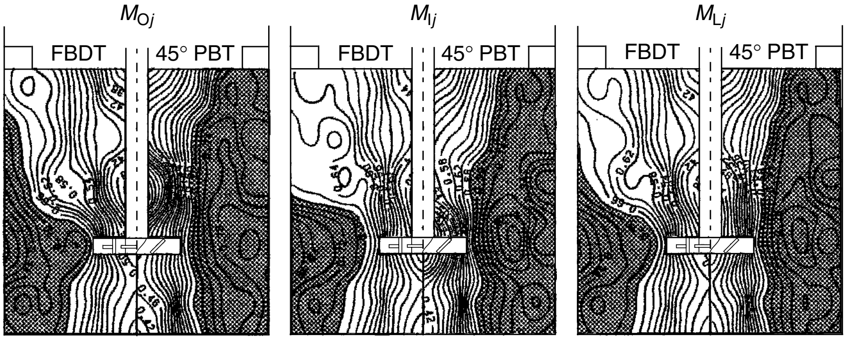


Figure 2.12 Local mixing capacity map in case of FBDT impeller and 45° PBT impeller in a stirred vessel (lines are contours of height at intervals of 0.02).

5. Noteworthy point

(a) In the case of the FBDT impeller, the indices for both the blender and the distributor take higher values in the impeller discharge flow region. On the other hand, in the case of the 45° PBT impeller, they take higher values at the intermediate part between the impeller axis and the vessel wall. The difference described above depends on the main flow produced by the impeller or, in other words, the radial flow by the FBDT impeller and the axial flow by the 45° PBT impeller affect the respective distributions of the local mixing capacity. (In the case of the 45° PBT impeller, the indices for both the distributor and the blender take smaller values at the impeller region.)

(b) It can be said that the impeller region in the case of the 45° FBT impeller only performs the function of a pipe to transport the fluid.

(c) The value of the whole mixing capacity of each impeller can be obtained by averaging all the local mixing capacities as follows:

FBDT impeller: 0.633
45° PBT impeller: 0.697

The 45° PBT impeller has a larger value than the FBDT impeller, and this result is contrary to the case in which the tracer is injected into the impeller position described in Challenge 2.3. This difference is caused by the difference in the local mixing capacity at the region where the impeller discharge flow reaches. In other words, the change in the spatial distribution of the tracer concentration with time is significantly affected by the position of the tracer injection.

6. Supplementary point

(a) Incidentally, the change in the spatial distribution of the tracer concentration with time for the case when the tracer is injected into a 10-region is

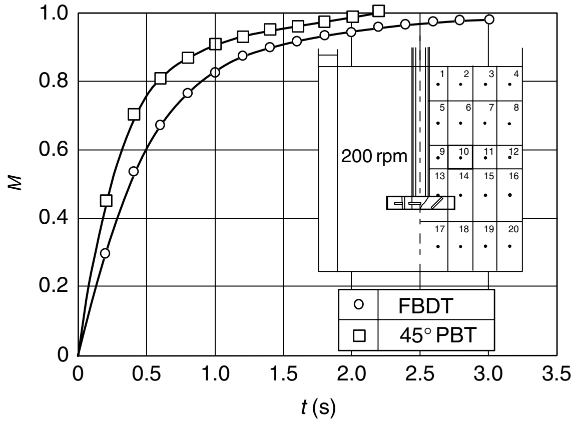


Figure 2.13 Mixedness change with time when tracer is injected from 10-region in case of FBDT impeller and 45° PBT impeller in a stirred vessel.

calculated based on the transition probability shown in Table 2.2. As shown in Figure 2.13, it is clear that the mixing rate of the 45° PBT impeller has a larger value than that of the FBDT impeller. This result is contrary to the case where the tracer is injected into the impeller position. This fact shows that the result based on the transient response method strongly depends on the position at which the tracer is injected.

(b) The evaluation of the mixing capacity based on the transition response method is significant only when the tracer injection position has a considerable role in the mixing operation, e.g., the mixing performance in a microorganism fomenter is discussed when the nutriment is fed from some specific position.

(c) It can be said that the evaluation based on the movement of an inner substance from one region to another region is indispensable for suitably discussing the mixing performance.

Challenge 2.8. Local mixing capacity of a circular pipe[10]

1. Scope

The difference in the spatial distribution of the tracer concentration between the cases of the injection into the pipe center and into the pipe wall ring, as shown in Challenge 2.5, can be considered to depend on the difference in the local mixing capacity among the radial positions. The mixing capacity of each radial position in a circular pipe has been discussed based on the distribution of the turbulent diffusivity in the horizontal cross-section of the pipe. However, investigating the mixing phenomena in a circular pipe based on the turbulent diffusivity is not sufficient, and it is necessary to discuss the local mixing capacity of each radial position.

2. Aim

To clarify the local mixing capacities $M_{\mathrm{O}j}$ and $M_{\mathrm{I}j}$ of each radial position in a circular pipe.

3. Experiment

(a) Apparatus

Circular pipe: circular pipe made of acrylic resign (inner diameter: 76 mm).
Device for three-dimensional measurement: mirror.

(b) Condition

Fluid: city water (room temperature).
Tracer: polystyrene spherical particle (0.8–1.2 mmϕ, density 1.0 g/cm^3, volume ratio of approximately 6.0×10^{-6}).
Flow state: fully developed turbulent flow ($Re = 1.1 \times 10^4$).
Test section: 25 cm in length at 150 cm downstream from the inlet.
Time interval to calculate the transition probability: 0.5 s.
Partition of the circular pipe: Figure 2.14(a) and (b) (square region: 3.6 mm × 3.6 mm × 10 mm).

(c) Tracer particle measurement method

Video camera.

(d) Procedure

After confirming that the flow state in a circular pipe becomes steady, the three-dimensional movement of water, which is represented by the movement of the tracer particles, is measured by recording their movement on video. The local mixing capacity M_{O} and $M_{\mathrm{I}j}$ based on the outflow and inflow, respectively, of each radial region as the distributor and blender, respectively, is calculated by using Eq. (2.27).

4. Experimental result

Figure 2.14(a) and (b) (Marks of tracer particles at the downflow region; (Dots represent particles).
Figure 2.14(c) (Local mixing capacities versus dimensionless radial position).

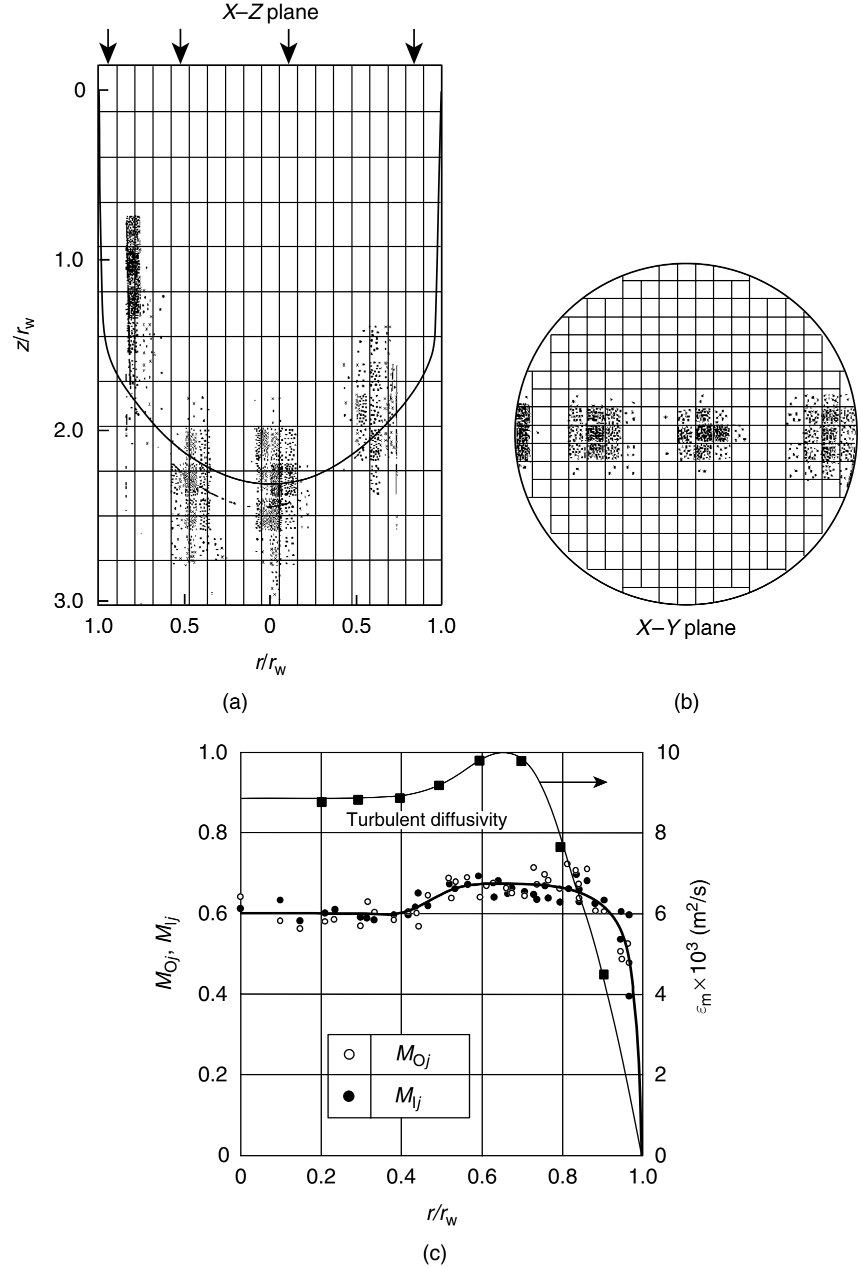


Figure 2.14 (a,b) The degree of dispersion of the tracer particles at downflow region is shown by number of dots. (a) Vertical cross-section, (b) Horizontal cross-section. (c) Local mixing capacities and turbulent diffusivity in a circular pipe.

5. Noteworthy point

(a) The radial distributions of the index as the blender and that as the distributor do not show a clear distribution, and both indices take a higher value in the region of $r/r_W = 0.5\text{–}0.9$.
(b) The result is very different from the radial distribution of the traditional turbulent diffusivity, as shown in Figure 2.14. This difference depends on the scale. The behavior of a polystyrene sphere depends on the turbulent velocity fluctuation that has a larger scale than the diameter of the styrene sphere (0.8–1.2 mm). On the other hand, the traditional turbulent diffusivity was measured after considering the turbulent velocity fluctuations that have a considerably smaller scale.
(c) The reason for the result obtained in Challenge 2.5 is that the pipe wall ring injection took a shorter distance than the pipe center injection for attaining a certain mixedness. Further, the difference lies not only between the size of the pipe wall ring area and the pipe center area but also between the local mixing capacity in both areas. The wall ring area has a higher mixing capacity than that in the center area. In other words, the tracer that is injected into the wall ring area immediately reaches the higher local mixing capacity region.

6. Supplementary point

(a) The experimental result described above is more significant when the focus of the study is macroscale turbulent fluctuations, rather than traditional turbulent diffusivity.

2.5 Evaluation of mixing performance of multi-component mixing

Until now, the evaluation of the mixing state and mixing capacity for the single-phase mixing system has been discussed. However, multi-component mixing is usually operated in practical industrial operations. In this section, the focus is on multi-component mixing in flow and batch systems[11].

(1) Multi-component mixing operation

As given in the following sections, when an element of the substance in the vessel is selected, the mixing state is evaluated from the viewpoint of information entropy based on the uncertainty regarding “the component of the element that is selected.” Additionally, a method of evaluation of the mixing capacity based on the change in the mixing state with time is discussed. The origin of time is set as the time when the impeller begins to revolve. In order to define the mixedness that quantitatively shows the degree of the mixing state of the multi-component system, the following conditions are set (Figure 2.15):

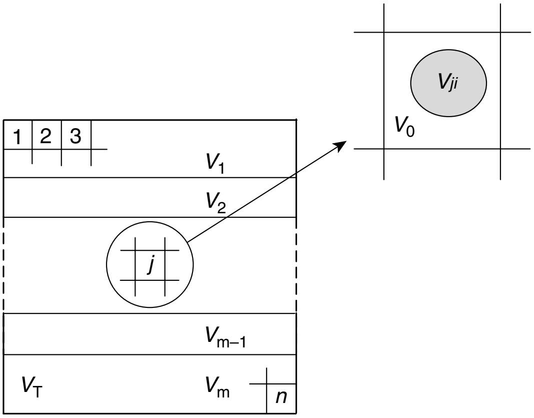


Figure 2.15 Definition diagram for multi-component mixing in a stirred vessel.

(1) The vessel (volume V_T) is partitioned into n imaginary regions with equal volume V_0:

$$nV_0 = V_T \tag{2.32}$$

(2) The volume of the i-component is V_i:

$$\sum V_i = \sum m_i V_0 = V_T \tag{2.33}$$

(3) The volume of the i-component in the j-region at time t is v_{ji}:

$$\sum v_{ji} = V_i \tag{2.34}$$

Under these conditions, the uncertainty regarding "the component of the element that is selected" is considered. Since the ratio of the volume of the i-component to the total volume of all components is V_i/V_T, this ratio is the probability that the selected element is the i-component. Therefore, the amount of information that is obtained by the news that gives the information that the selected element is the i-component is expressed as

$$I(C_i) = -\log \frac{V_i}{V_T}$$

The probability that the news described above is given is V_i/V_T. Therefore, the amount of uncertainty regarding "the component of the element that is selected" is expressed by the information entropy as

$$H(C) = \sum_i^m \frac{V_i}{V_T} I(C_i) = -\sum_i^m \frac{V_i}{V_T} \log \frac{V_i}{V_T} \equiv -\sum_i^m P_i \log P_i \tag{2.35}$$

In practice, it is natural that the region into which each component is fed is already known. In such a case, some relationship between each component and each region after the beginning of the mixing operation still remains. Therefore, the amount of uncertainty described above (Eq. (2.35)) is expected to decrease if the position from where the element is selected is known. The uncertainty regarding "the component of the element that is selected when the element is taken from the j-region" is discussed.

Since the i-component volume in the j-region is v_{ji}, the ratio v_{ji}/V_0 becomes the probability that the selected element in the j-region is the i-component. Therefore, the amount of information that is obtained by the news that informs that the selected element is the i-component is expressed as

$$I(C_i/j) = -\log \frac{v_{ji}}{V_0}$$

The probability that the news described above is given is v_{ji}/V_0. Therefore, the amount of uncertainty regarding "the component of the element that is selected when the element is taken from the j-region" is expressed by the information entropy as

$$H(C/j) = -\sum_{i}^{m} \frac{v_{ji}}{V_0} I(C_i/j) = \sum_{i}^{m} \frac{v_{ji}}{V_0} \log \frac{v_{ji}}{V_0} \equiv \sum_{i}^{m} P_{ji} \log P_{ji} \tag{2.36}$$

It is not decided that the element is always selected from the j-region and the probability of this event is V_0/V_T. The amount of uncertainty regarding "the component of the element that is selected" is expressed by the conditional entropy as

$$H(C/R) = \sum_{j}^{n} \frac{V_0}{V_T} H(C/j) = -\frac{1}{n} \sum_{j}^{n} \sum_{i}^{m} P_{ji} \log P_{ji} \tag{2.37}$$

In other words, by obtaining the news that the selected region is reported, the amount of uncertainty regarding "the component of the element that is selected" decreases from $H(C)$ to $H(C/R)$. The decrease in the amount of information is mutual entropy, which is written as

$$\begin{aligned} I(C;R) &= H(C) - H(C/R) \\ &= -\sum_{i}^{m} P_i \log P_i + \frac{1}{n} \sum_{j}^{n} \sum_{i}^{m} P_{ji} \log P_{ji} \end{aligned} \tag{2.38}$$

This is the amount of information that is obtained by the news that the region from where the element is selected is known. When no mixing occurs and each region is occupied by only one component that is identical to the beginning of the operation, the component of the selected element is determined by reporting the news that the selected region is reported and the uncertainty regarding "the component of the element that is selected" disappears. Therefore, in this case, the mutual entropy $I(C; R)$ has a value identical to the self-entropy $H(C)$. When perfect mixing occurs and each region is occupied by every component in the same volume ratio as the ratio of each component volume to the total volume of all components, the news that the selected region is reported has no value; further, the same amount of uncertainty remains even after obtaining the news. Therefore, in this case, the mutual entropy $I(C; R)$ has zero value. This result coincides with the human experience.

Next, the probability function P_{ij} for the maximum and minimum values of $I(C; R)$ is discussed mathematically. The self-entropy $H(C)$ in Eq. (2.38) is decided only by the fraction of each component in the feed, and the value does not change through the mixing process. Then, the maximum and minimum values of the mutual information entropy are determined by the value of the conditional entropy $H(C/R)$. Since the range of the variable j is fixed as $\underline{1 \le j \le n}$, according to Section 1.7, $H(C/R)$ takes the minimum and maximum values as

$$H(C/R)_{\min} = 0 \quad \text{at} \quad P_{ij,i=a} = 1, \quad P_{ij,i\neq a} = 0 \tag{2.39a}$$

$$H(C/R)_{\max} = -\sum P_i \log P_i \quad \text{at} \quad P_{ij} = V_i/V_T = P_i \tag{2.39b}$$

where "a" is some specific component.

Therefore, the mutual entropy $I(C; R)$ has the minimum and maximum values as

$$I(C; R)_{\max} = -\sum P_i \log P_i \quad \text{at} \quad P_{ij,i=a} = 1, \quad P_{ij,i\neq a} = 0 \tag{2.40a}$$

$$I(C; R)_{\min} = 0_i \qquad \text{at} \quad P_{ij} = V_i/V_T = P_i \tag{2.40b}$$

where "a" is some specific component.

Whether the conditions for the maximum and minimum values of $I(C; R)$ correspond with the practical mixing phenomena is considered. The condition under which $I(C; R)$ has the maximum value is realized when the substance in a region does not disperse to any other region or flow out to another region without dispersion. In other words, the condition for the latter case occurs when the piston flow is established. On the other hand, the condition under which the mutual entropy assumes the minimum value is realized when the substance in a region flows out equally to all regions. In other words, this condition occurs when perfect mixing flow is established. Therefore, the multi-component mixedness can be defined as the degree of approach from the no-mixing state to the perfect-mixing state by using the maximum and minimum values of $I(C;\, R)$ as follows:

$$M(m) = \frac{I(C;\, R)_{\max} - I(C;\, R)}{I(C;\, R)_{\max} - I(C;\, R)_{\min}} = \frac{-(1/n)\sum_{j}^{n}\sum_{i}^{m} P_{ji}\log P_{ji}}{-\sum_{i}^{m} P_i \log P_i} \tag{2.41}$$

The mixedness defined by Eq. (2.41) takes a value from zero for the no-mixing state to unity for the perfect-mixing state.

$$0 \le M(m) \le 1 \tag{2.42}$$

In practice, it is not easy to measure the volume of each component at each small region in the equipment. In such a case, in general, the equipment is partitioned into suitable volume regions in which the mixing can be assumed to be perfect. In this case, by considering that each partitioned region is a collection of small equivolume regions described above, the same treatment described above is possible. In other words, there is no difference in the result between the case where the vessel is partitioned into suitable volume regions and that where the vessel is partitioned into small equivolume regions, although the method of summation should be available.

Until now, the quantitative evaluation of the multi-component mixedness based on the spatial distribution of each component in the vessel was discussed. It is obvious that the evaluation of the mixing capacity should be performed based on the change in the multi-component mixedness with time. However, few investigations exist on the multi-component mixing capacity based on the change in mixedness with time.

Next, the case of $m = n$ in Eq. (2.41), that is, the case that the number of components equals the number of imaginary equivolume partitioned regions, is considered. In this case, the denominator in Eq. (2.41) is log n and the formula coincides with the definition of the whole mixing capacity defined by Eq. (2.30). This is the point of contact between the multi-component mixedness and whole mixing capacity; that is, both indices need not be discussed separately.

All the indices discussed until now are shown in Table 3.2, and the quantitative relationships among the indices are clarified. When one component in the multi-component mixing is treated as the tracer, that is, when the components other than the tracer component are treated without discrimination, the mixedness based on the state of the tracer dispersion has the following relationship with the multi-component mixedness:

$$M(m) = \frac{-\sum_{i}^{m} M_i P_i \log P_i}{-\sum_{i}^{m} P_i \log P_i} \tag{2.43}$$

This fact implies that the mixedness by using the transient response method is a special case of the multi-component mixedness[12].

Challenge 2.9. Change in mixedness with time for FBDT impeller and 45° PBT impeller in the case of five-component mixing

1. Scope

Although multi-component mixing is typically operated in practical industrial operations, the number of investigations regarding the mixing performance in the multi-component mixing is very small. In particular, there have been few investigations on the difference in multi-component mixing capacity among the impeller styles.

2. Aim

To clarify the difference in the mixing capacity M defined by Eq. (2.43) for five-component mixing between the FBDT impeller and 45° PBT impeller.

3. Calculation

(a) Apparatus

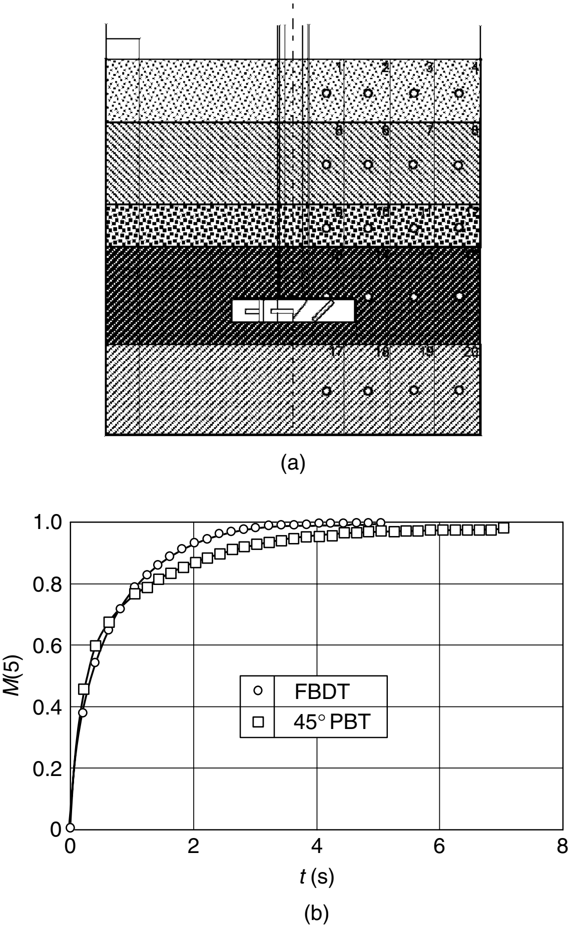
Stirred vessel: Figure 2.7 (Cylindrical flat bottom vessel, four baffles).
Impeller: Figure 2.7(b) (FBDT impeller and 45° PBT impeller).

(b) Condition

Fluid: ion exchange water.
Impeller rotational speed (N): 200 rpm ($Re = 3.61 \times 10^4$).
Time interval: 0.2 s.
Partition of stirred vessel: Figure 2.7(a).
Initial setting of component: Figure 2.16(a) (utilize the partition).

(c) Method

The transition probabilities P_{ij} and P_{ji} in 0.2 s at an impeller rotational speed of 200 rpm for each impeller in Table 2.2 are used for estimating



(a)

(b)

Figure 2.16 (a) Initial setting of five components in a stirred vessel. (b) Mixedness change with time of five-component mixing in case of FBDT impeller and 45° PBT impeller in a stirred vessel.

the change in the spatial distributions of the concentration of each component with time, and the mixedness *M* defined by Eq. (2.43) is calculated.

4. Calculated result

Figure 2.16(b) (Five-component mixedness versus time *t*).

5. Noteworthy point

(a) In the first stage, the 45° PBT impeller has a higher value of mixedness than that of the FBDT impeller. In the next stage, however, the FBDT impeller has a higher value of mixedness than that of the 45° PBT impeller. This difference depends on the main flow produced by the impeller and the initial setting of the five components in the vessel. In other words, in the first stage, the radial flow produced by the FBDT impeller flows through the same component layer, and *M* takes a smaller value. On the other hand, the axial flow produced by the 45° PBT impeller flows vertically through the different component layers in the first stage, and *M* takes a larger value.

(2) Mixedness of multi-phase mixing

Multi-phase mixing is often seen in industries. In general, the distribution of not only the dispersed phase but also the continuous phase depends on the local position in the equipment in the case of a multi-phase operation such as gas–liquid mixing system, liquid–liquid mixing system, solid–liquid mixing system, and gas–liquid–solid mixing system. In order to evaluate the mixing state in such systems, both the dispersed phase and continuous phase should be considered.

It is possible to define an evaluation index for the mixing state by using the definition of multi-component mixedness in the previous section. The following discussion focuses on the mixing state of the continuous phase and the dispersed phase with a particle size distribution.

The distribution of the dispersed particle size is divided into $m - 1$ groups in the order of size, and each group is considered to be individual component. Additionally, the continuous phase is treated as another component. From this consideration, the mixing can be treated as m-component mixing, and the multi-component mixedness defined by Eq. (2.43) in the previous section can be applied. The extended definition of mixedness for the mixing of the continuous phase and dispersion phase can be expressed as

$$M(m) = \frac{-\sum_{j}^{n} (V_j/V_T)\left(p_{jC}\log p_{jC} + \sum_{j}^{m-1} p_{ji}\log p_{ji}\right)}{-P_C \log P_C - \sum_{j}^{m-1} P_i \log P_i} \tag{2.44}$$

where

$$P_{jC} = \frac{v_{jC}}{V_j}, \quad p_{ji} = \frac{v_{ji}}{V_j}, \; P_C = \frac{\sum_j^n v_{jC}}{V_T}, \; P_i = \frac{V_i}{V_T} \tag{2.45}$$

and V_i is the volume of the i-component; V_j the volume of the j-region; v_{ji} the volume of the i-component in the j-region; v_{jC} the volume of the continuous phase in the j-region; and V_T the total volume of the equipment. Additionally, the volume V_0 should be smaller than the total volume of any component. The denominator of Eq. (2.44) is the self-entropy for the uncertainty regarding "the component of the element that is selected" when an element is selected from the equipment. The numerator of the equation is the mutual entropy obtained from the news that the position where the element is selected is known. The mixedness defined above assumes a value from zero for the case that each region is occupied by only one component to unity for the case that each region is occupied by all the components in the ratio of $V_{i/}V_T$:

$$0 \leq M(m) \leq 1 \tag{2.46}$$

However, the mixing of the dispersed and continuous phases is considered here, and it is possible to apply the same way of thinking for plural dispersed phases. Additionally, the newly defined mixedness can be applied to judge whether the assumption of MSMPR (mixed suspension mixed product removal) in the crystallization operation is established.

Let us deal with MSMPR

This type of equipment is sometimes called the circulating-magma crystallizer. In this type, the uniformity of suspension of the product solids within the crystallizer is sufficient. It has been said that most commercial equipment satisfies this assumption.

Challenge 2.10. Suitability of assumption of MSMPR in a crystallizer

1. Scope

The aim of crystallization is to separate the observed component into higher quality crystals. The crystal size and probability density distribution of its size become very important factors for the product or the following processes. Although multi-phase mixing is fairly common in industries, there have been few investigations on the mixing performance of operations/equipment. In crystallization operation, the assumption of MSMPR has been used to design a crystallizer without a detailed discussion. Therefore, the assumption of MSMPR must be studied quantitatively.

2. Aim

To clarify the suitability of the assumption of MSMPR by calculating the mixedness M defined by Eq. (2.44) as a typical example.

3. Calculation

(a) Apparatus

Crystallizer: Figure 2.17(a) (Cylindrical flat bottom vessel).

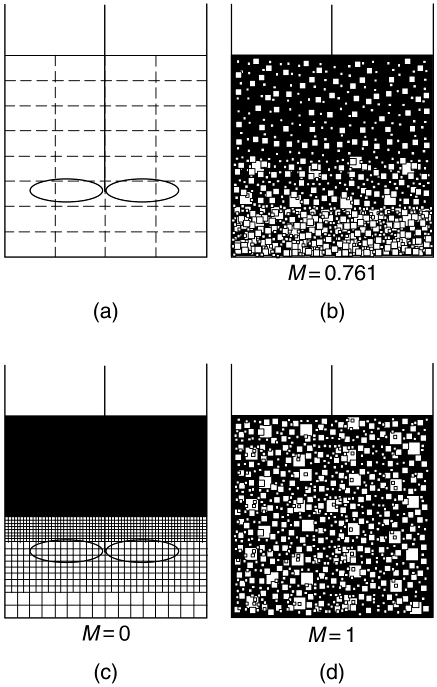


Figure 2.17 Mixing in a crystallizer considering continuous phase (upper left-hand: imaginary regions partitioned; the others: local size distribution of dispersion particle and mixedness).

(b) Condition

Particle size distribution: Figure 2.17(b)–(d) (three groups; volume ratio of one particle—16:4:1).

Volume ratio of three components and continuous liquid phase:

Table 2.3 (1:2:1:4, homogeneous in the radial direction).

Partition of stirred vessel: Figure 2.17(a) (32 concentric doughnut-type regions).

Table 2.3 Volume fraction of all components.

	Continuous phase	Disperse Phase		
		Largest Size Crystals	Medium Size Crystals	Smallest Size Crystals
Volume fraction (%)	50	12.5	25	12.5

(c) Method

From the volume ratio of the three particle groups and continuous liquid phase, the mixedness M defined by Eq. (2.44) is calculated.

4. Calculated result

Figure 2.17(b) (General case).
Figure 2.17(c) (Each region is occupied by one component).
Figure 2.17(d) (Each component is equally occupied in each region).

5. Noteworthy point

(a) The MSMPR assumption must be carefully applied. (In the practical crystallization, the density of the crystal is generally larger than that of the continuous liquid phase, and the crystal has a tendency to settle.)

Challenge 2.11. Relationship between mixedness and impeller rotational speed in particle–particle–liquid mixing

1. Scope

There are four aims of solid particle–liquid mixing:

(1) to obtain a homogeneous slurry,
(2) to prevent the sedimentation of a solid particle,
(3) to control mass transfer or reaction between a solid particle and a liquid,
(4) to control the crystal size in crystallization.

Many operations treat particle–liquid mixing in chemical industry. The first aim of solid–liquid mixing is to make a solid particle float. However, mixing performance of operations/equipment is not clear. Additionally, when many kinds of particles are involved, it is not known whether there is any difference in the mixedness between the following two cases: the case where all particles are treated as a particle (two-phase mixing—particle and continuous liquid phase) and the case where every particle is treated individually (multiple-phase mixing—each particle and continuous liquid phase). Therefore, a solution to this unsolved problem is not imperative.

2. Aim

To clarify the difference in mixedness M defined by Eq. (2.44) between the following two cases—the case where the two kinds of particles are treated as a particle (two phase mixing; particle and continuous liquid phase) and the case where each particle is treated individually (three-phase mixing; each particle and continuous liquid phase).

3. Experiment

(a) Apparatus

Stirred vessel: Figure 2.18 (Cylindrical flat bottom vessel, four baffles).
Impeller: FBDT impeller.

(b) Condition

Fluid: ion exchange water.
Impeller rotational speed (N): 40–700 rpm ($Re = 0.722 \times 10^4$–12.6×10^4).

Particle:

(i) Glass sphere (average diameter 401 μm (351–451 μm), 592 μm(491–701 μm), specific gravity 2.5, volume ratio 1–2.5%).
(ii) Ion exchange resin sphere (average diameter 395 μm (330–460 μm), 625 μm (500–750 μm), specific gravity 1.21, volume ratio 1–2.5%).

(c) Procedure

The stirred vessel that involves fixed solid particles and ion exchange water is placed in a square water tank. After the flow state in the stirred vessel becomes steady under a fixed impeller rotational speed, images

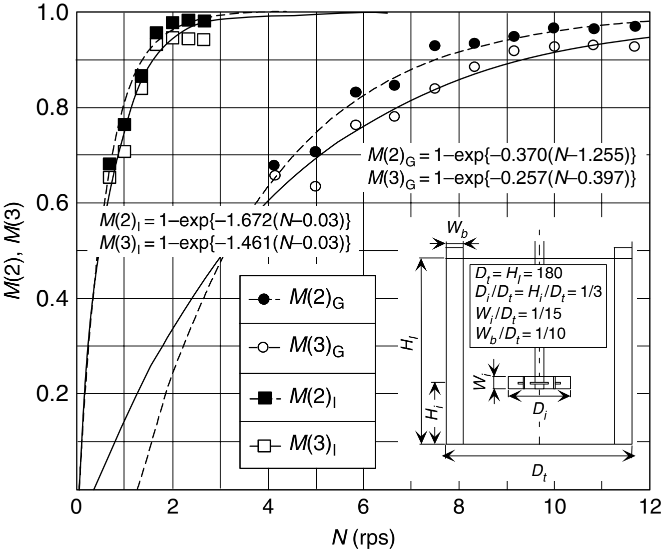


Figure 2.18 Solid–liquid mixing in a stirred vessel.

of the particles in the vessel are taken by irradiating through a 4 mm slit. Based on the image, the concentration probability density distribution of each phase is calculated and the mixedness M defined by Eq. (2.44) is calculated.

4. Experimental result

Figure 2.18 (Mixednesses of two-phase mixing versus the impeller rotational speed for the two cases.)

5. Noteworthy point

(a) $M(2)_{\mathrm{G}}$ clearly takes a larger value than $M(3)_{\mathrm{G}}$. However, in the case of an ion exchange resin sphere, the difference described above is negligibly small.
(b) The changes in $M(2)_I$ and $M(3)_I$ with N are greater than that in the case of a glass sphere. This is because the specific gravity of the glass sphere is large, and although the glass sphere can move on the bottom of the vessel, it cannot float. In other words, the area of movement of the glass sphere is too narrow.
(c) The change in $M(2)_I$ and $M(2)_{\mathrm{G}}$ with N has larger values than that of $M(3)_I$ and $M(3)_{\mathrm{G}}$.
(d) The change in $M(2)_{\mathrm{G}}$, $M(3)_{\mathrm{G}}$, $M(2)_I$, and $M(3)_I$ with N can be expressed as

$$M(2)_{\mathrm{G}} = 1 - \exp\{-0.370(N - 1.255)\}$$
$$M(3)_{\mathrm{G}} = 1 - \exp\{-0.257(N - 0.397)\}$$
$$M(2)_{\mathrm{I}} = 1 - \exp\{-1.672(N - 0.03)\}$$
$$M(3)_{\mathrm{I}} = 1 - \exp\{-1.461(N - 0.03)\}$$

6. Supplementary point

(a) Since the curves of $M(2)_{\mathrm{G}}$, $M(3)_{\mathrm{G}}$, $M(2)_I$, and $M(3)_I$ intersect the horizontal axis (impeller rotational speed), the point of intersection can be defined as the mixing start impeller rotational speed. From a practical viewpoint, this is significant information in the case that the specific gravity of a particle is greater than the continuous liquid phase such as a glass sphere.

2.6 Summary

In this chapter, the following points have been clarified:

(1) The following evaluation indices of mixing operations/equipment are defined on the basis of information entropy. The evaluation indices have clear quantitative relationships.

For a flow system:

Mixedness based on the residence time probability density distribution.

For a batch system:

Mixedness based on the transient response method.
Local mixing capacity based on the transient probability of the inner substance from one region to another in the vessel.
Whole mixing capacity based on the average value of the local mixing capacity.
Mixedness for multi-component mixing based on the distribution of each component in the vessel.
Mixedness for multi-phase mixing based on the distribution of each phase in the vessel.

(2) Several examples of the application of newly defined indices of the mixing performance to mixing operations/equipment are shown; further, their usefulness is clarified:

   (a) Relationship between the mixing capacity and the number of tanks in completely mixing equivolume tanks in a series model (continuous stirred tank reactors model) in the case of a flow-stirred vessel
   (b) Relationship between the mixing capacity and impeller rotational speed and setting positions of the inlet and outlet in a flow-stirred vessel;
   (c) Relationship between the mixedness and impeller rotational speed in an aerated stirred vessel
   (d) Relationship between the change in mixedness with distance in the axial direction and tracer injection position in the radial direction in a circular pipe
   (e) Relationship between the mixedness and gas flow rate in a bubble column
   (f) Local and whole mixing capacities of the FBDT-type impeller and 45° PBT-type impeller
   (g) Local mixing capacity of a circular pipe
   (h) Change in mixedness with time of the FBDT-type impeller and 45° PBT-type impeller in the case of five-component mixing
   (i) Suitability of the assumption of MSMPR in a crystallizer
   (j) Relationship between the mixedness and impeller rotational speed in particle–particle–liquid mixing

(3) In order to evaluate the mixing capacity of batch mixing operation/equipment absolutely, the transient probability of an inner substance from one region to another region should be considered because the result of the transient response method depends on the region where the tracer is injected.

CHAPTER 3

Separation Phenomena

3.1 Introduction

The separation operation/equipment and mixing operation/equipment are related to each other as the front and rear of an event.

In Chapter 2, the evaluation indices of mixing operations/equipment are discussed. Therefore, as a logical continuation from the preceding chapter, it is natural to discuss the evaluation of separation operations/equipment in this chapter.

In general, in industrial processes, the necessity of a separation operation occurs after the mixing/reaction operations. Therefore, the separation operation is an indispensable process for chemical engineering. The purpose of the separation process is different from that of the mixing process, and the purpose of separation is only to separate substances depending on the requirements. However, identical to the mixing phenomenon, the separation phenomenon is definitely a function of space and time. The separation equipment is classified into two groups:

(1) batch system,
(2) flow system.

Let us deal with separation operations in chemical engineering. The representative separation operations are extraction, absorption, and crystallization.

The aim of extraction is to promote mass transfer or extraction reaction between two phases (a mutually insoluble liquid–liquid system) by dispersing one liquid phase in another. In principle, creating a larger interface area in this operation is advantageous. However, it is necessary to consider that interface phenomena depend on the type of system. Extractors are classified into three types:

(1) tower type extractor (perforated plate tower, packed tower, baffle tower, and so on),
(2) mixer–settler type extractor,
(3) centrifugal type extractor.

The aim of absorption is to purify gas, recover useful material, remove harmful material, and so on. In absorption, a solute or multiple solutes are dissolved in liquid by contacting gas and liquid (absorbent). Absorption is classified into two groups:

(1) physical absorption,
(2) chemical absorption.

Further, the absorption equipment is classified into two groups:

(1) gas dispersion type (bubble column, aerated stirred vessel, plate/tray column, etc.),
(2) liquid dispersion type (packed column, spray tower, wetted-wall column, etc.).

The most important parameter for the equipment is the gas–liquid contact area per unit volume, and the driving force for mass transfer is sufficiently large.

The aim of crystallization is not only to separate the solute by purification but also to produce particles. In general, by cooling the solution or vaporizing the solvent, crystals of high purity are produced and separated from the mother liquor. The method based on super-saturated concentration is classified into five groups:

(1) cool type (exchange heat),
(2) vacuum type (reduce pressure to vaporize the solvent),
(3) reaction type (create sediment),
(4) lowering solubility type (add a poor solvent to reduce the solubility),
(5) pressure type (add pressure to increase saturated temperature). However, it is not easy to obtain crystals of a particular particle size probability density distribution. Recently, the demand for crystals of a higher purity and more homogeneous size has increased.

It is not an exaggeration to say that the evaluation of mechanical separation process/equipment is performed based on the conception of yield and quality. For example, the widely used Newton efficiency is defined in the following manner. In the case of better separator, both the yield and the degree of quality improvement should have higher values. From another viewpoint, the recovery rate of useful component should have a higher value, and the intermix rate of the useless component into product should have lower values for the better separator. In other words, this definition is based on the concept that for a better separator, either the product of hopeful factors has a higher value or the difference in the value between the hopeful factor and the hopeless factor has larger values. Then, in the definition of Newton efficiency, the product

$$(\text{yield}) \times (\text{improvement of quality})$$

or the balance

$$(\text{recovery rate of useful component}) - (\text{intermix rate of useless component into product})$$

is used. Newton efficiency is then defined as the ratio of the values for a practical case and an ideal case. In the binary component system (Figure 3.1, Table 3.1), Newton efficiency is expressed as

$$\eta_N = \frac{P}{F}\frac{x_P P}{x_F F} = \frac{x_P P}{x_F F} - \frac{(1-x_P)P}{(1-x_F)F} = \frac{(x_P - x_F)(x_F - x_R)}{x_F(1-x_F)(x_P - x_R)} \tag{3.1}$$

There is Richarse' efficiency as another representative separation efficiency. This Richarse' efficiency depends on the product

$$(\text{recovery rate of useful component in product}) \times (\text{recovery rate of useless component in residuum})$$

This definition is based on the consideration that the recovery rate of both the useful and useless components should have a higher value for a good separator.

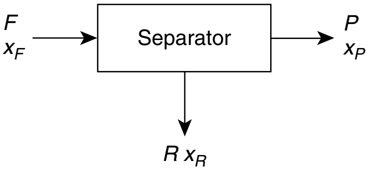


Figure 3.1 Feed, product, and residuum in case of binary component in a separation equipment.

Table 3.1 Separation of binary component.

	Flow Rate	Fraction of Useful Component
Feed	F	x_F
Product	P	x_P
Residuum	R	x_R

This line of thinking is almost the same as that for Newton efficiency. Richarse' efficiency in this case is expressed as

$$\eta_R = \frac{x_P P}{x_F F}\frac{(1-x_R)R}{(1-x_F)F} = \frac{(1-x_R)(x_P-x_F)(x_F-x_R)x_P}{(1-x_F)(x_P-x_R)^2 x_F} \tag{3.2}$$

As observed, it is impossible to require the relevancy to the intuitive definition of separation efficiency and traditional mixing performance indices that are shown in Table 2.1. In this chapter, a new separation efficiency will be defined by making use of the information entropy on the basis that there is no discrepancy between the viewpoint of the definitions of the evaluation indices for mixing and separation operations/equipment.

3.2 Definition of separation efficiency[1]

Here, the method of evaluating the separation state for an operation condition is discussed based on the consideration that the separation efficiency and mixedness should be related as the front and rear of an event. As given now, when an element of a substance in the vessel is selected, the separation state is evaluated from the viewpoint of information entropy based on the uncertainty regarding "the component of the element that is selected." In order to define the separation efficiency that shows the degree of separation state quantitatively, the following conditions are set (Figure 3.2):

(1) The vessel (volume V_T) is partitioned into n imaginary regions with equal volume V_0:

$$nV_0 = V_T \tag{3.3}$$

(2) The volume of the i-component volume is V_i:

$$\sum_i V_i = \sum_i m_i V_0 = V_T \tag{3.4}$$

(3) The volume of the i-component in the j-region at time t is v_{ji}:

$$\sum_j v_{ji} = V_i \tag{3.5}$$

Besides, it is possible to apply the same consideration in the following discussion to the case of a flow system when the region is replaced by an outlet. (The discussion from this point until Eq. (3.11) is identical to that from Eq. (2.32) to Eq. (2.40) in the previous Chapter.)

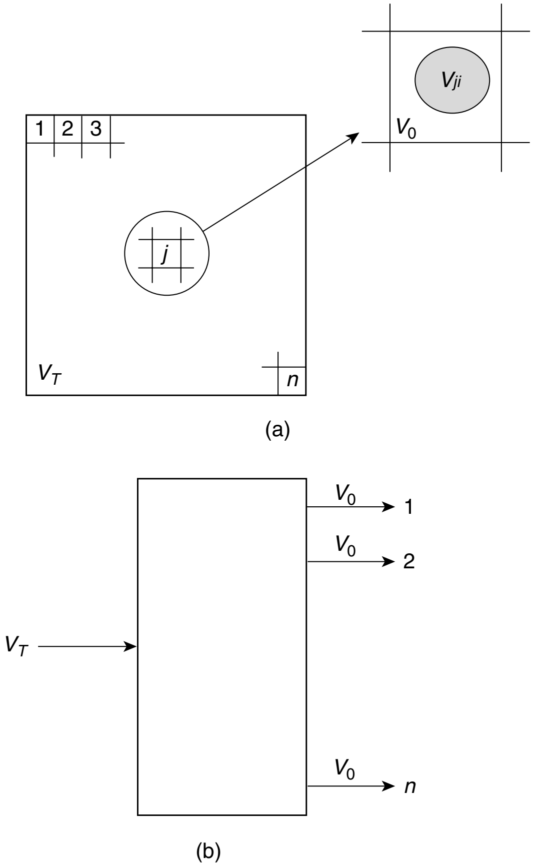


Figure 3.2 Definition diagram for separation process.

Under these conditions, the uncertainty regarding "the component of the element that is selected" is considered. Since the ratio of the volume of the i-component to the total volume of all components is V_i/V_T, this ratio becomes the probability that the selected element is the i-component. Therefore, the amount of information that is obtained by the news that gives the information that the selected element is the i-component is expressed as

$$I(C_i) = -\log \frac{V_i}{V_T}$$

The probability that the news described above is given is V_i/V_T. Therefore, the amount of uncertainty regarding "the component of the element that is selected" is expressed by the information entropy as

$$H(C) = \sum_i^m \frac{V_i}{V_T} I(C_i) = -\sum_i^m \frac{V_i}{V_T} \log \frac{V_i}{V_T} \equiv -\sum_i^m P_i \log P_i \tag{3.6}$$

In practical operations, it is natural that the region into which each component is fed is already known. There is also a relationship between each component and each region after the beginning of the separation operation. Therefore, the amount of uncertainty described above (Eq. (3.6)) should decrease if the position from where the element is selected is known. Next, the uncertainty regarding "the component of the element that is selected when the element is taken from the j-region" is discussed.

Since the i-component volume in the j-region is v_{ji}, the ratio v_{ji}/V_0 becomes the probability that the selected element in the j-region is the i-component. Therefore, the amount of information that is obtained by the news which informs that the selected element is the i-component is expressed as

$$I(C_i/j) = -\log \frac{v_{ji}}{V_0}$$

The probability that the news described above is given is v_{ji}/V_0. Therefore, the amount of uncertainty regarding "the component of the element that is selected when the element is taken from the j-region" is expressed by the information entropy as

$$H(C/j) = -\sum_i^m \frac{v_{ji}}{V_0} I(C_i/j) = \sum_i^m \frac{v_{ji}}{V_0} \log \frac{v_{ji}}{V_0} \equiv \sum_i^m P_{ji} \log P_{ji} \tag{3.7}$$

It is not decided that the element is always selected from the j-region and the probability of this event is V_0/V_T. The amount of uncertainty regarding "the component of the element that is selected" is expressed by the conditional entropy as

$$H(C/R) = \sum_j^n \frac{V_0}{V_T} H(C/j) = -\frac{1}{n} \sum_j^n \sum_i^m P_{ji} \log P_{ji} \tag{3.8}$$

In other words, by obtaining the news that the selected region is reported, the amount of uncertainty regarding "the component of the element that is selected" decreases from $H(C)$ to $H(C/R)$. The decrease in the amount of information is the mutual entropy, which is written as

$$\begin{aligned} I(C;R) &= H(C) - H(C/R) \\ &= -\sum_i^m P_i \log P_i + \frac{1}{n} \sum_j^n \sum_i^m P_{ji} \log P_{ji} \end{aligned} \tag{3.9}$$

This is the amount of information that is obtained from the news that the region from where the element is selected is known. When perfect separation occurs and each region is occupied by only one component similar to the beginning of the operation, the component of the selected element is determined by reporting the news that the selected region is reported; and the uncertainty regarding “the component of the element that is selected” disappears. Therefore, in this case, the mutual entropy $I(C; R)$ has the same value as that of self-entropy $H(C)$. When no separation occurs and each region is occupied by every component in the same volume ratio as the ratio of each component volume to the total volume of all components, the news that the selected region is reported has no value; further, the same amount of uncertainty remains after obtaining the news. Therefore, in this case, the mutual entropy $I(C; R)$ becomes zero. This result is identical to observations from human experience.

Next, the probability function P_{ij} for the maximum and minimum values of $I(C; R)$ is discussed mathematically. The self-entropy $H(C)$ in Eq. (3.9) is determined only by the fraction of each component in the feed, and the mixing process does not change the value. Then, the maximum and minimum values of the mutual information entropy are determined by the value of the conditional entropy $H(C/R)$. Since the range of variable j is fixed as $1 \le j \le n$ according to Section 1.7, $H(C/R)$ has the following minimum and maximum values:

$$H(C/R)_{\min} = 0 \quad \text{at} \quad P_{ij,i=a} = 1, \quad P_{ij,i\neq a} = 0 \tag{3.10a}$$

$$H(C/R)_{\max} = -\sum P_i \log P_i \quad \text{at} \quad P_{ij} = V_i/V_T = P_i \tag{3.10b}$$

where “a” is some specific component.

Therefore, the mutual entropy $H(C; R)$ has the following minimum and maximum values:

$$I(C; R)_{\max} = -\sum P_i \log P_i \quad \text{at} \quad P_{ij,i=a} = 1, \quad P_{ij,i\neq a} = 0 \tag{3.11a}$$

$$I(C; R)_{\min} = 0_i \quad \text{at} \quad P_{ij} = V_i/V_T = P_i \tag{3.11b}$$

where “a” is some specific component.

Under these conditions, the uncertainty regarding “the component of the element that is selected” is considered. In fact, the following discussion is the same as that in the definition of the multi-component mixedness. Only the practical conditions corresponding to the maximum and minimum values of the mutual entropy are different from the case of the definition of the multi-component mixedness. According to Section 1.7, the mutual entropy $I(C; R)$ has the following minimum and maximum values:

$$I(C; R)_{\max} = -\sum P_i \log P_i \quad \text{at} \quad P_{ij,i=a} = 1, \quad P_{ij,i\neq a} = 0 \tag{3.12a}$$

$$I(C; R)_{\min} = 0 \quad \text{at} \quad P_{ij} = V_i/V_T = P_i \tag{3.12b}$$

where “a” is some specific component. The condition under which $I(C; R)$ has a maximum value is realized when perfect separation is established, that is, when each region is occupied by one expected component. The condition under which $I(C;R)$ has the minimum value is established when no separation occurs, that is, when each region is occupied by all the components in a ratio equal to that of the feed. Therefore, the separation efficiency can be defined as the degree of approach from the no separation state to the perfect separation state by using the maximum and minimum values of $I(C;R)$ as follows:

$$\eta(m) = \frac{I(C;R)_{\min} - I(C;R)}{I(C;R)_{\min} - I(C;R)_{\max}} = 1 - \frac{-(1/n)\sum_{j}^{n}\sum_{i}^{m} P_{ji}\log P_{ji}}{-\sum_{i}^{m} P_i \log P_i} \tag{3.13}$$

The index defined by Eq. (3.13) has a value from zero for the no separation state to unity for the perfect separation state:

$$0 \le \eta(m) \le 0 \tag{3.14}$$

In practice, it is not easy to measure the volume of each component at each small region in the equipment. Then, in general, the equipment is partitioned into suitable volume regions in which the mixing can be assumed to be perfect. In this case, by considering that each partitioned region is the collection of small equivolume regions described above, the same treatment is possible. In other words, there is no difference in the result between the case that the vessel is partitioned into suitable volume regions and the case that the vessel is partitioned into small equivolume regions, though the way of summation should be available.

Until now, it has been possible to quantitatively evaluate the separation efficiency based on the spatial distribution of each component in the vessel. It is obvious that the separation capacity should be evaluated based on the change in separation efficiency with time. However, investigations on the separation capacity based on the change in separation efficiency with time are few. This fact depends on the flow separation system being the main current and the batch separation system being the secondary current.

The difference between this separation efficiency and that from the multi-component mixedness is that the correspondence of the conditions for the maximum and minimum values of the mutual entropy is reversed. Therefore, it is natural that this separation efficiency has the following clear relationship with the multi-component mixedness:

$$M(m) + \eta(m) = 1 \tag{3.15}$$

Finally, it can be said that a consistent method of evaluating the mixing performance and separation performance is established. That is, the first aim of this book is attained sufficiently. The relations among the indices of mixing and separation are shown in Table 3.2.

Table 3.2 Quantitative relationship among indices of mixing and separation performance.

	Mixedness for m component $M(m)$ and Separation Efficiency $\eta(m)$		
	$M(m) = \dfrac{-\sum_j^n \sum_i^m \frac{1}{n} p_{ji} \log p_{ji}}{-\sum_i^m p_i \log p_i} \Rightarrow M(m) + \eta(m) = 1$		
	$\Downarrow m = n$		
	Blender	distributor	mixedness for impulse response method
Whole	$M(n) = M_{\mathrm{IW}} = \dfrac{-\sum_j^n \sum_i^n \frac{1}{n} p_{ij} \log p_{ij}}{\log n}$	$= M_{\mathrm{OW}} = \dfrac{-\sum_j^n \sum_i^n \frac{1}{n} p_{ji} \log p_{ji}}{\log n}$	$\Rightarrow i = \mathrm{O} \Rightarrow$ $M_{\mathrm{IW}} = \dfrac{-\sum_j^n p_{j\mathrm{O}} \log p_{j\mathrm{O}}}{\log n}$
	$\Uparrow \sum_j$	$\Uparrow \sum_j$	
	Blender	distribution	
Region-j	$M_{\mathrm{I}j} = \dfrac{-\sum_j^n p_{ji} \log p_{ji}}{\log n}$	$M_{\mathrm{O}j} = \dfrac{-\sum_j^n p_{ij} \log p_{ij}}{\log n}$	

Challenge 3.1. Binary component separation (comparison with Newton efficiency)

1. Scope

Newton efficiency is the most widely used separation efficiency. As discussions regarding the capability of detection of Newton efficiency are insufficient, it is necessary to clarify the capability of detection of Newton efficiency. Additionally, a new separation efficiency is defined by Eq. (3.13) based on the information entropy. Consequently, it is necessary to compare the capability of the detection of newly defined separation efficiency with that of Newton efficiency.

2. Aim

To compare the capability of detection under several operation conditions between the new separation efficiency defined by Eq. (3.13) and Newton efficiency. (To compare the change in the newly defined separation efficiency with the value of the fraction of the useful component in the feed, product, and residuum with that of Newton efficiency.)

3. Calculation

(a) Condition

Component: binary component. (The binary component separation operation is the simplest example for discussing the separation efficiency.)

Value of the fraction of useful component: full range. (The same result is obtained by reversing the viewpoints with regard to not only the useful and useless components but also the product and the residuum.)

(b) Method

By consistently changing the value of the fraction of the useful component in the feed, product, and residuum, the separation efficiencies defined by Eq. (3.13) and Newton efficiency defined by Eq. (3.1) are calculated.

4. Calculated result

Figure 3.3(a) (New separation efficiency and Newton efficiency versus the fraction of the useful component in the residuum under each operation condition. The same result is obtained by reversing the viewpoints with respect to not only the useful and useless components but also the product and residuum.)

Figure 3.3(b) (Dimensionless new separation efficiency versus dimensionless fraction of useful component. The value on the vertical axis is the dimensionless separation efficiency obtained by using the maximum value of the separation efficiency under each operation condition.

5. Noteworthy point

(a) When x_R value is close to x_F, Newton efficiency shows a high response to the change in x_R. However, when x_R is close to 0, Newton efficiency shows a very weak response to the change of x_R value.

(b) On the other hand, the newly defined separation efficiency shows an almost constant response for the change in x_R across all ranges of x_R.

(c) The newly defined separation efficiency is more useful than Newton efficiency; in particular, it has an advantage that high-grade separation operation is performed.

Challenge 3.2. Evaluation of separation performance of distillation column

1. Scope

The distillation column is a typical equipment in chemical industries. Traditionally, the separation performance of the distillation column is evaluated on the basis of the composition of distillate and residue; and separation efficiency such as Newton

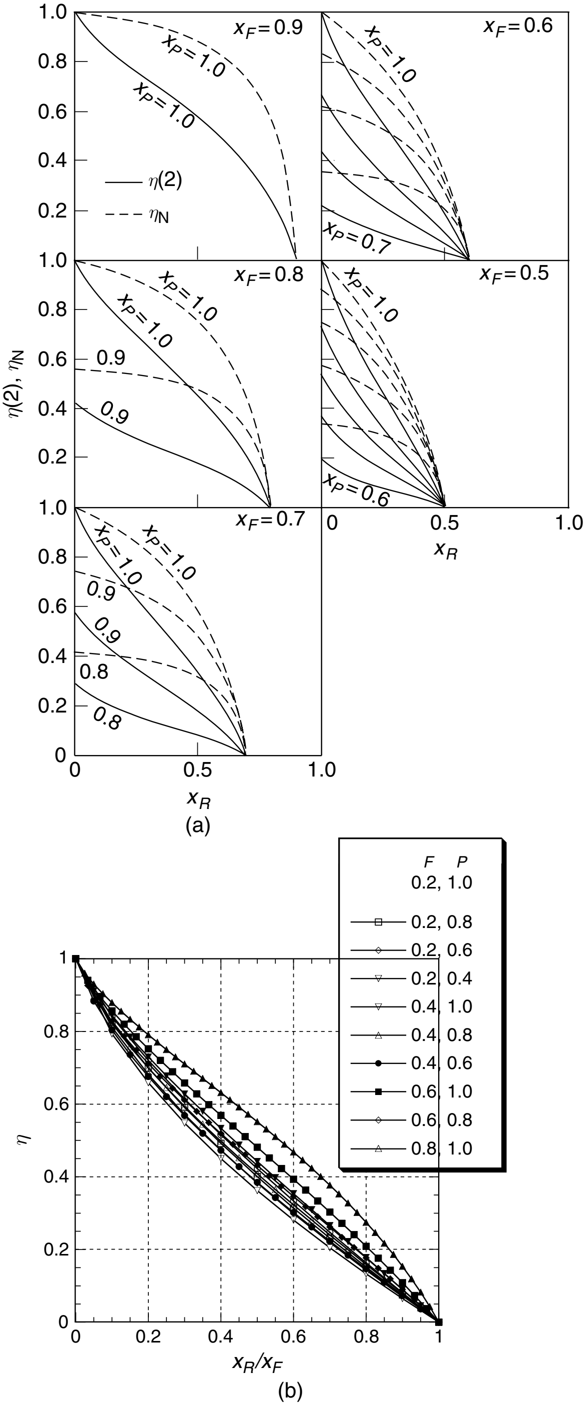


Figure 3.3 **(a) Comparison of new efficiency curves and Newton efficiency curves. (b) S-shaped curve of new efficiency.**

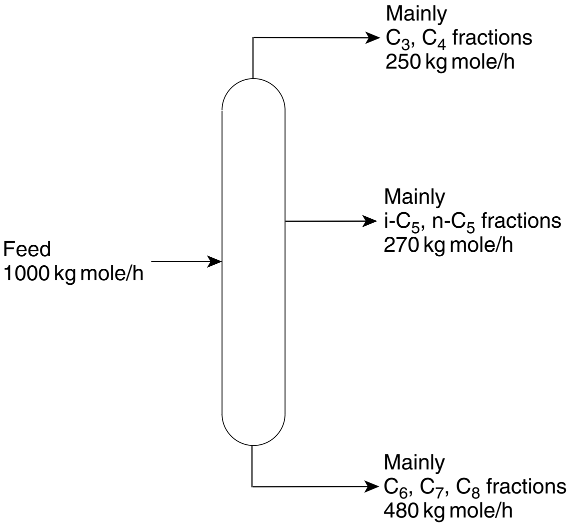


Figure 3.4 Distilation column.

efficiency has seldom been used. However, the component that must be separated and its flowout exit are fixed at the design stage of the distillation column. Therefore, it is possible to apply the new separation efficiency defined by Eq. (3.13) to evaluate the separation performance of the distillation column; such an investigation might develop new steps in the design of a distillation column.

2. Aim

To apply the separation efficiency defined by Eq. (3.13) to the evaluation of the separation performance of the distillation column.

3. Calculation

(a) Apparatus: Figure 3.4 (Typical distillation column as an LPG; liquefied petroleum gas separator (C_3 and C_4) from pentanes and heavy component mixture).
(b) Condition

Feed: Table 3.3 (1000 kg mole/h of the mixture of a heavy component and pentanes (250 kg mole/h of C_3 and C_4, 270 kg mole/h of C_5, 480 kg mole/h of C_6, C_7, and C_8)).

Theoretical output: Table 3.3 (C_3 and C_4 from top; isoC_5 and normal C_5 from middle of column; C_6, C_7, and C_8 from bottom).

Actual output: Table 3.3 (isoC_5 and normal C_5 except C_3 and C_4 from top; C_3, C_4, C_6, C_7, and C_8 except isoC_5 and normal-C_5 from middle; isoC_5 and normal-C_5 except C_6, C_7, and C_8 from bottom).

Number of components: Table 3.3 (top component (C_3 and C_4); middle component (C_5); and bottom component (C_6, C_7, and C_8)).

Table 3.3 Initial conditions and the sensitivity of new separation efficiency in case of distillation operation.

		Feed (1000 kg mole/h)			$\eta(3)$
		C_3, C_4 (250 kg mole/l)	C_5 (270 kg mole/l)	C_6, C_7, C_8 (480 kg mole/h)	
Case I	C_3, C_4	240	10	0	
	C_5	10	250	10	0.805
	C_6, C_7, C_8	0	0	470	
Case II	C_3, C_4	230	20	0	
	C_5	20	230	20	0.721
	C_6, C_7, C_8	0	20	460	

(c) Method

Under the condition described in Table 3.3, the new separation efficiency defined by Eq. (3.13) is calculated.

4. Calculated result

Table 3.3 (New separation efficiency and operation conditions).

5. Noteworthy point

(a) It is efficient and appropriate to evaluate the separation performance of the distillation column by the new separation efficiency defined by Eq. (3.13). As shown in Table 3.3 (from Case I to Case II), the new separation efficiency responds sensitively to a small change in the composition of the distillate and residue (e.g., when the ratio of the intermixture increases by a few per cent; (top component becomes the middle component, the bottom component becomes the intermediate component, and the middle component becomes the top and bottom components); hence, the new separation efficiency decreases by 12%. This fact shows the new separation efficiency has a very high detection sensitivity.

3.3 Summary

The following points have been clarified in this chapter:

(1) The evaluation index of the separation operation/equipment is defined by using information entropy. This newly defined index has a clear quantitative relationship with the mixedness that was defined in Chapter 2.

(2) The newly defined separation efficiency is applied to the following examples, and the usefulness of the index is clarified:

  (a) binary component separation (comparison with Newton efficiency),
  (b) evaluation of the separation performance of the distillation column.

(3) The newly defined separation efficiency shows reasonable values and an almost constant response to the change in the value of the fraction of the useful component in feed or product or residuum.

CHAPTER 4

Turbulent Phenomena

4.1 Introduction

Fluid flow usually affects many phenomena that occur in chemical equipment. Keeping in mind that almost all the operations in chemical equipment are performed in a turbulent flow field and few operations are performed in a laminar flow field, knowledge on the structure of a turbulent flow field is very important for a discussion regarding the phenomena in the chemical equipment.

Turbulent fluid motion is defined as follows: "An irregular flow condition in which various quantities such as velocity and temperature show a random variation with space and time coordinates, and a statistically distinct value is discerned." Experiments have yielded most of the knowledge on turbulent flows.

Turbulence is classified into two groups based on the cause of occurrence:

(1) wall turbulence (by frictional forces at fixed walls (flow through conduits, flow past bodies)),
(2) free turbulence (by shear stress between fluid layers).

Further, turbulence is also classified into two groups based on the characteristics of randomness:

(1) pseudo-turbulence (with a regular pattern that exhibits a distinct constant periodicity in time and space (Kármán vortex)),
(2) real turbulence (exhibits a random variation with time and space coordinates (general shear turbulence)).

Additionally, the ideal turbulence for investigation is classified into three groups:

(1) homogeneous turbulence (turbulent statistical values do not change by parallel movement of coordinates),
(2) isotropic turbulence (turbulent statistical values do not change by rotation and reflection of coordinates),
(3) homogeneous isotropic turbulence (turbulent statistical values do not change by parallel movement, rotation, and reflection of coordinates).

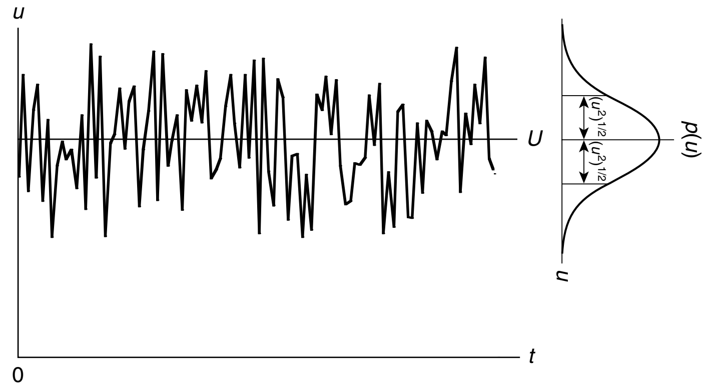


Figure 4.1 Velocity fluctuations with time.

In general, the physical values at one point in the turbulent flow field are treated by dividing them into time average value and fluctuating value. For example, when the variation in the velocity at a point u is measured with respect to time t, as shown in Figure 4.1, the time average velocity with respect to time T is defined as

$$U = \lim_{T\to\infty} \frac{1}{T}\int_0^T u \, \mathrm{d}t$$

Therefore, the momentary velocity u at a point can be expressed as the sum of the time average velocity U and fluctuating velocity u':

$$u = U + u'$$

Let us deal with the equation of motion for turbulent flow. In the case of laminar flow under the condition of constant density and constant viscosity, the equation of motion is expressed by the Navier–Stokes equation as

$$\rho\left(\frac{\partial u_i}{\partial t} + u_j \frac{\partial u_i}{\partial x_j}\right) = -\frac{\partial P}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_j}\right)$$

In the case of turbulent flow under the condition of constant density and constant viscosity, the equation of motion can be obtained by replacing the velocity u with $U + u'$ and so on. By considering an averaging process, the equation of motion for turbulent flow under the condition of constant density and constant viscosity is obtained as

$$\rho\left(\frac{\partial \overline{U_i}}{\partial t} + \overline{U_j}\frac{\partial \overline{U_i}}{\partial x_j}\right) = -\frac{\partial \overline{P}}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial \overline{U_i}}{\partial x_j} - \rho \overline{u_i' u_j'}\right)$$

The difference between this equation for turbulent flow and the Navier–Stokes equation for laminar flow is the Reynolds stress/turbulent stress term $-\rho\overline{u_i'u_j'}$ appears in the equation of motion for turbulent flow. This equation of motion for turbulent flow involves non-linear terms, and it is impossible to be solved analytically. In order to solve the equation in the same way as the Navier–Stokes equation, the Reynolds stress or fluctuating velocity must be known or calculated. Two methods have been adopted to avoid this problem—phenomenological method and statistical method. In the phenomenological method, the Reynolds stress is considered to be proportional to the average velocity gradient and the proportional coefficient is considered to be turbulent viscosity or mixing length:

$$-\rho\overline{u_i'u_j'} = -\rho\varepsilon_{ij}\left(\frac{\partial\overline{U_i}}{\partial x_j} + \frac{\partial\overline{U_j}}{\partial x_i}\right)$$

$$-\rho\overline{u_i'u_j'} = -\rho\ell^2\left|\frac{\partial\overline{U_i}}{\partial x_j}\right|\frac{\partial\overline{U_i}}{\partial x_j}$$

where ε_{ij} is the turbulent viscosity and ℓ the mixing length. In the former case, it is considered that the viscosity increases by changing from a laminar flow μ to a turbulent flow ε_{ij}. From this, the equation of motion for turbulent flow becomes the same level as that of the Navier–Stokes equation for laminar flow. However, turbulent viscosity is not a physical property, and the value of turbulent viscosity depends on the flow condition. In the latter case, the mixing length corresponds to the mean free path of gas movement in the kinetic theory of gases. However, this mixing length depends on the flow condition and is not constant. We now focus on the mechanism of turbulent flow. In a chemical equipment, the flow changes continuously because of instability. The overlapping of turbulent motion into motions on various length scales is useful because the different scales play somewhat different roles in the dynamics of motion. Therefore, this turbulent motion is often expressed as the motion of eddies of different sizes. A turbulent eddy is very useful for the development of turbulence descriptions. Small eddies contribute to larger wave number components of the spectrum; the spectrum curve is often loosely interpreted in terms of the energy associated with eddies of various sizes. It is true that the essential characteristics of the turbulent structure that strongly affects transport phenomena, such as the mutual relation among eddies, are not clarified even by solving the equation of motion of turbulent flow by using the value of Reynolds stress. In such a case, statistical method can be introduced. By considering that the motion of fluid at a point is affected by the motion of the surrounding fluid, the motion of a fluid at two points is investigated by using the statistical theory, and the Kármán–Howarth equation is derived as

$$\frac{\partial}{\partial t}(u'^2 f) - u'^3 \frac{1}{r^4}\frac{\partial}{\partial r}(r^4 k) = 2\nu u'^2 \frac{1}{r^4}\frac{\partial}{\partial r}\left(r^4 \frac{\partial f}{\partial r}\right)$$

where f is the coefficient of special longitudinal velocity correlation; k the coefficient of special triple velocity correlation; r the distance between two points; u' the root mean square turbulent velocities.

This Kármán–Howarth equation involves a double correlation that corresponds to the Reynolds stress.

Further, an interesting question is how the kinetic energy of turbulence will be distributed according to various eddies/frequencies. Such a distribution of the energy among the eddies/frequencies is usually termed the energy spectrum. Our focus is now on the double correlation in the Kármán–Howarth equation, and finally, the dynamic equation for the energy spectrum that is obtained by the Fourier transform of the double correlation is derived as

$$\frac{\partial}{\partial t}E(k, t) = F(k, t) - 2\nu k^2 E(k, t)$$

where $E(k, t)$ is the three-dimensional energy spectrum function; $F(k, t)$ the three-dimensional transfer spectrum function; and k the wave number.

By comparing the above dynamic equation and experimental results, the concept of cascade process is presented. In the cascade process, turbulent energy is supplied from main flow and the energy is transmitted from larger eddies to smaller eddies sequentially. The dissipation energy through the process takes larger value at smaller eddies and there is the smallest eddy that is not broken up any more. There is a statistical lower limit to the size of the smallest eddy; there is a minimum scale of turbulence that corresponds to a maximum frequency in the turbulent motion. The lower limit of eddies is determined by the viscosity effect and the value decreases with increasing average velocity. The upper size limit of eddies is determined by the size of the apparatus. However, the energy spectrum function that shows turbulent structure is not yet clear. For a chemical engineer who considers the scale-up of equipment is established by building up the same turbulent structure, the statistical theory is significant.

A method of discovering time scales associated with turbulent motion is Fourier analysis. Along with the frequency spectrum, the wave number spectra can be defined through the Fourier transforms of space correlations. (When applicable, Taylor's hypothesis can be used to derive a spatial spectrum from an observed time spectrum.) The wave number is an important parameter in many theoretical treatments of turbulent motion.

$E(k)$ indicates the distribution of energy across different length scales, where k is the magnitude of the wave number:

$$u'^2 = \int_0^\infty E(k)\mathrm{d}k$$

When the expression of $E(k)$ in an equipment is the same as that in another equipment, it can be said that the structure of the turbulent flow fields in both equipment is identical. Anyway, turbulent phenomena can be expressed as probability terms. In other words, a discussion of the structure of turbulence in a chemical equipment is possible by making use of information entropy.

Hereupon, let us move to the relationship among frequency f, wave number k, and size of eddy ℓ.

In general, the turbulent structure is discussed based on the following relationships among the above three:

$$k = \frac{2\pi f}{U}$$

$$\ell \propto \frac{1}{k}$$

where U is the average velocity.

4.2 Probability density distribution function for velocity fluctuation

As mentioned above, the characteristics of turbulent flow is defined as "various quantities show a random variation with space and time coordinates." Figure 4.1 shows an example of change in velocity with time, that is, the velocity fluctuations in a fully developed turbulent flow. The intensity of velocity fluctuation is constant and can be expressed as

$$u'^2 = \frac{1}{T}\int_0^T \{u(t) - U\}^2\mathrm{d}t \tag{4.1}$$

This intensity is the variance in the probability density distribution of the velocity fluctuation. The velocity fluctuation that has the characteristics described above can be discussed from the viewpoint of information entropy as follows. The probability density distribution that takes the maximum value of information entropy under the condition that the value of variance is fixed should be a normal distribution, as mentioned in Section 1.7. Therefore, the probability density distribution of the velocity fluctuation would be a normal distribution if the velocity behaves in such a manner that the information entropy takes the maximum value. In practice, the probability density distribution of the velocity fluctuation in a fully developed turbulent flow shows a normal distribution, as shown in Figure 4.1. This shows that the velocity fluctuation in a fully developed turbulent flow occurs as the information entropy attains the maximum value. The other fluctuation of the physical quantity in a fully developed turbulent flow can be understood in the same way as described above.

4.3 Energy spectrum probability density distribution function

There have been many investigations on turbulent flow. Most of these investigations have paid attention to the treatment of Reynolds stress; this is because it is impossible to obtain the analytical solution of a momentum equation of turbulent flow for the existence of non-linear terms such as the Reynolds stress term. The Reynolds stress is a double correlation and has a strong relation with the energy spectrum probability density distribution (ESD) function. In the case of the turbulent flow structure, ESD function is discussed based on the extended line of the dynamic equation. ESD shows the degree of contribution of each eddy/wave number to the velocity fluctuations, that is, the weight of each scale fluctuation on the turbulent kinetic energy. However, the formula expressing ESD is not clarified, and a few relations with the wave number are presented for limited wave number ranges, as shown in Table 4.1[1–8]. ESD is a probability density distribution function that has a continuous variable of the wave number.

Table 4.1 Traditional **ESD** *Function.*

Wave number range	Low	Medium	Higher	Highest
$E(k) \propto$	k (Chandrasekhar[1], Rotta[2], Prudman[3]) k^2 (Birkhoff[4]) k^4 (Loitsansky[5])	k (Ogawa[6])	$k^{-5/3}$ (Kolmogoroff[7])	k^{-7} (Heisenberg[8])

In the following section, the turbulent flow structure is discussed based on ESD by using information entropy.

Let us deal with the energy spectrum.

Spectrum analysis is a random data analysis method that probably has origins in the periodgraph that was obtained on the basis of the period of the variation in the sunspot 150 years ago by the English physicist Arthur Schster.

The velocity fluctuation in a turbulent flow is the synthesis of many different frequency waves, and Fourier integral and Fourier transform are two of the mathematical expressions of the structure. When $u'(t)$ is a real fluctuation, the following relation is obtained:

$$u'(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)\mathrm{e}^{j\omega t}\mathrm{d}\omega = \int_{-\infty}^{\infty} F(\omega)\mathrm{e}^{j\omega t}\mathrm{d}f \qquad (\omega = 2\pi f)$$

Energy content E is considered, and the following relation is obtained based on Parseval's theorem:

$$E = \int_{-\infty}^{\infty} |u'(t)|^2\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2\mathrm{d}\omega = \int_{-\infty}^{\infty} |F(\omega)|^2\mathrm{d}f$$

The intensity $|F(\omega)|^2$ is called the energy spectrum of $u'(t)$ or energy spectral density function of $u'(t)$ and is interpreted as the contribution from frequency $f(= kU/2\pi)$ to the turbulence energy. Along with the frequency spectrum, the wave number spectra can be defined through the Fourier transform of space correlations. (When applicable, Taylor's hypothesis can be used to derive a special spectrum from an observed time spectrum.) If the periodic function or the value in the limited region becomes zero without focusing on the region, the value of $|F(\omega)|^2$ also becomes limited. However, if the region that is focused is infinity, the average energy per unit time is calculated and referred to as the power spectrum:

$$S(f) = \lim_{T\to\infty}\left[\frac{1}{T}|F(f)|^2\right]$$

In general, the energy spectrum is calculated by using the auto-correlation function $R_{11}(\tau)$ based on Wiener–Khintchine's theorem as follows:

$$|F(\omega)|^2 = \int_{-\infty}^{\infty} R_{11}(\tau)\mathrm{e}^{-j\omega\tau}\mathrm{d}\tau = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u(t)u(t-\tau)\mathrm{e}^{-j\omega\tau}\mathrm{d}\tau$$

This $|F(\omega)|^2$ corresponds to $E(k)$ because $\omega = 2\pi f$ and $k = 2\pi f/U$.

The turbulent flow field can be considered to be a non-linear system, and the turbulent flow structure that consists of various size eddies is affected by the non-linear terms. The formula that expresses ESD is discussed based on the uncertainty regarding "the wave number of the fluctuation that is selected" when a fluctuation is selected from the turbulent flow field[9–13]. In order to define the ESD formula, the following conditions and assumptions are set:

(1) The turbulent flow field consists of the basic eddy group and its sub-harmonic eddy groups that are generated sequentially from the basic eddy group.
(2) Each eddy group has an average scale (average wave number or average frequency), and the following relationship exists between the average wave number of the i-eddy group and that of the next $i+1$-eddy group.

$$K_{i+1}/K_i = 1/\alpha$$

(3) ESD of each eddy group is the one that gives the maximum amount of information entropy.
(4) The following relational equation exists between the turbulent kinetic energy of the i-eddy group and that of the $i+1$-eddy group.

$$P_{i+1}/P_i = 1/\beta$$

The collection of eddies that appear for identical reasons is called eddy group in this section. When the cascade process of turbulent kinetic energy transfer is considered, it can be understood that the smallest eddy group, that is, the basic eddy group, is determined naturally for an individual fluid, according to its property of viscosity.

Under these conditions, the uncertainty regarding "the wave number of the fluctuation that is selected" is discussed. By considering that ESD of the i-eddy group, $E_i(k)/u_i^2$ (k: wave number, u_i^2: kinetic energy of the i-eddy group) is a probability density distribution function and should satisfy the second assumption described above; the standardized conditions can be written as (hereafter, u' is referred to as u in order to simplify the discussion):

$$\int_0^\infty \frac{E_i(k)}{u_i^2}\mathrm{d}k = 1 \tag{4.2}$$

$$\int_0^\infty k\frac{E_i(k)}{u_i^2}\mathrm{d}k = K_i \tag{4.3}$$

The information entropy for the uncertainty regarding "the wave number of the fluctuation that is selected from the i-eddy group" is expressed by the information entropy according to Eq. (1.3) as

$$H_i(k) = -\int_0^\infty \frac{E_i(k)}{u_i^2}\log\frac{E_i(k)}{u_i^2} \tag{4.4}$$

The $E_i(k)/u^2$ function for the maximum value of the information entropy under the conditions of Eqs. (4.2) and (4.3) is obtained as follows according to Section 1.7:

$$\frac{E_i(k)}{u_i^2} = \frac{1}{K_i}\exp\left(-\frac{k}{K_i}\right) \tag{4.5}$$

Therefore, the function described above shows ESD for the *i*-eddy group.

Since the ratio of the turbulent kinetic energy of the *i*-eddy group to the total turbulent kinetic energy is P_i, the entire ESD can be obtained by adding the entire eddy group's ESD as

$$\begin{aligned}\frac{E(k)}{u^2} &= \sum_i^m \frac{P_i}{K_i}\exp\left(-\frac{k}{K_i}\right)\\ &= \frac{1}{K_1\sum\limits_j^m (1/\beta)^{j-1}}\sum_j^m\left\{\left(\frac{\alpha}{\beta}\right)^{j-1}\exp\left(-\alpha^{j-1}\frac{k}{K_1}\right)\right\}\end{aligned} \tag{4.6}$$

where K_1 is the average wave number of the basic eddy group.

If the α and β values are determined, ESD can be fixed. However, it is difficult to determine these values theoretically. In such a case, searching the most suitable values of α and β by plotting ESD curves practically for many combinations of α and β values is inevitable. Some examples of ESD for some representative combinations of α and β values are shown in Figure 4.2 in which the horizontal axis is k/K_1, the vertical axis is $E(k)K_1/u^2$, and the value of K_1 is fixed as unity.

The most suitable combination of α and β values is decided by comparing the result with the traditional knowledge on ESD as follows:

(1) There exists a clear wave number range in which Kolmogorov's $-5/3$ law can be applied.
(2) ESD decreases monotonously with an increase in the wave number and does not exhibit any fluctuations in this behavior.

As a result, the most suitable combination of the α and β values is determined as

$$\alpha = 3 \quad \text{and} \quad \beta = 1/2$$

When this combination of α and β values is used, the wave number range in which Kolmogorov's $-5/3$ law can be applied increases with an increase in the number of eddy groups. The entire ESD with the combination of the values of $\alpha = 3$ and $\beta = 1/2$ can be written as

$$\frac{E(k)}{u^2} = \frac{1}{K_1\sum\limits_j^m 2^{j-1}}\sum_j^m\left\{6^{j-1}\exp\left(-3^{j-1}\frac{k}{K_1}\right)\right\} \tag{4.7}$$

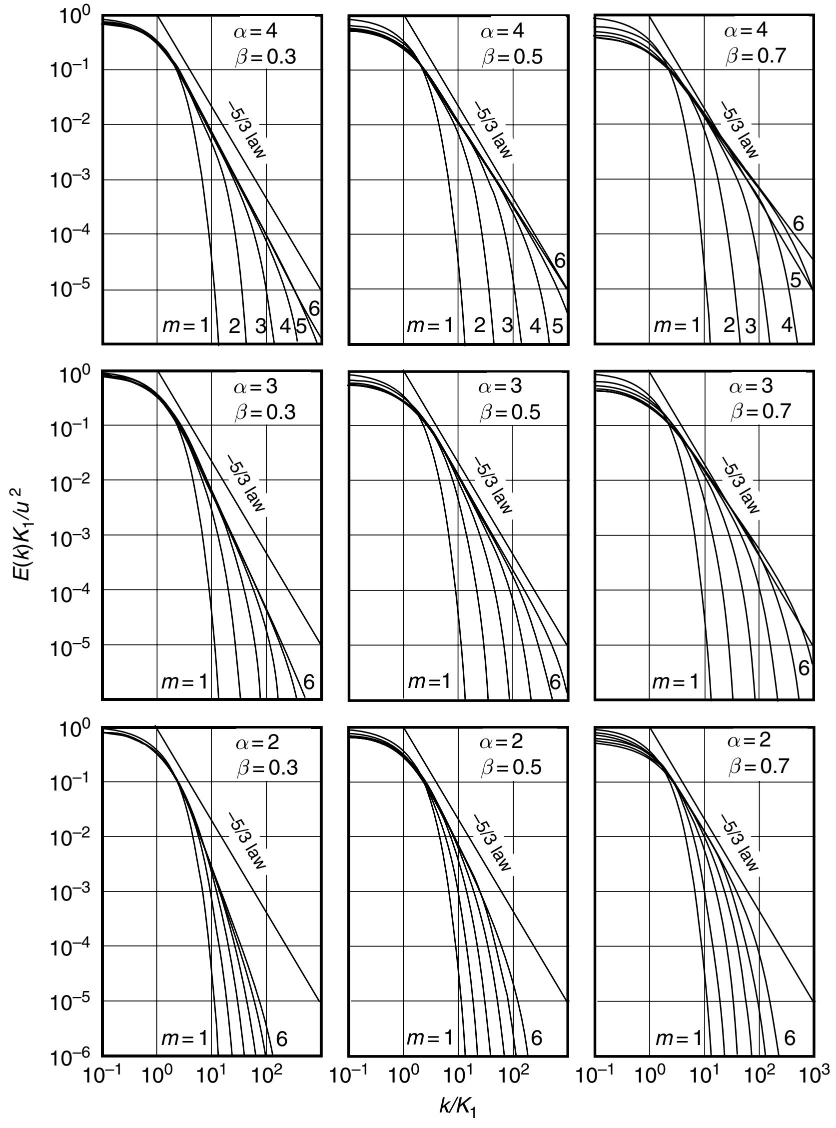


Figure 4.2 Effect of combination of values of α and β on ESD.

This result shows that the turbulent flow field consists of the basic eddy group and its sub-harmonic eddy groups that are generated sequentially from the basic eddy group according to the rule that the wave number and kinetic energy of the $i+1$-eddy group are 1/3 times and 2 times those of the i-eddy group, respectively. The fact that the ratio of the average wave number is 1/3 corresponds to the fact that the ratio of the frequency of the sub-harmonic group is 1/3 in most non-linear systems.

Challenge 4.1. Application of newly defined **ESD** *function to measured results*

1. Scope

Many experimental results of ESD have been reported and there are no common characteristics among them. Therefore, it is indispensable to examine to what extent the new ESD defined by Eq. (4.7) can be applied to the measured results.

2. Aim

To clarify the applicability of the new ESD defined by Eq. (4.7) to the measured results of ESD.

3. Calculation

(a) Condition

Data quoted: Figure 4.3 (Air flow in a circular pipe, air jet flow, and air and water downstream of a grid).

(b) Method

The experimental results of ESD are compared with the fitted theoretical ESD curves based on Eq. (4.7) by changing the value of m.

4. Calculated results

Figure 4.3 (ESD versus wavenumber)

5. Noteworthy point

(a) Each experimental result is sufficiently placed on a theoretical curve, and it can be said that the experimental result can be expressed by the new ESD defined by Eq. (4.7) for any value of m.

4.4 Scale of turbulence and turbulent diffusion

A discussion on the scale of turbulence and turbulent diffusion that can be derived from the newly defined ESD is given as follows.

ESD expressed by Eq. (4.7) can be rewritten by using the relational equation of $k = 2\pi f/U$ (U: average velocity (m/s); f: frequency (1/s)) as

$$E(f) = \sum_{i}^{m} \frac{P_i u^2}{F_i} \exp\left(-\frac{f}{F_i}\right) \tag{4.8}$$

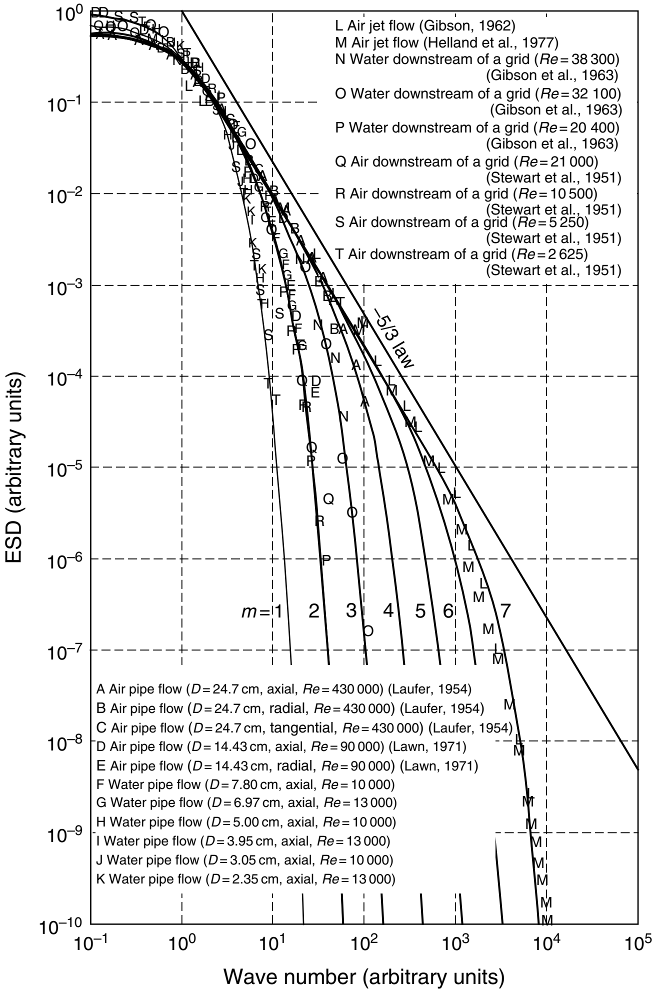


Figure 4.3 Estimated curves based on new ESD function and practical data of ESD[12,14–17].

where F_i is the average frequency. When this equation is Fourier transformed by using the theorem of Wiener–Khintchine, the following equation with regard to the double correlation is obtained:

$$R(t) = \frac{1}{u^2}\int_0^{\infty} E(f)\cos(2\pi f t)\mathrm{d}n = \sum_i^m \frac{P_i}{1+(2\pi F_i t)^2} \tag{4.9}$$

For $t \to 0$, this double correlation can be expressed by a parabolic curve as

$$R(t) \cong 1-(2\pi)^2 \sum_i^m P_i F_i^2 t^2 \tag{4.10}$$

According to the expression of $R(t)$ described above, the micro time and macro time scales can be determined as

$$\tau_0 = \frac{1}{2\pi\left(\sum_i^m P_i F_i^2\right)^{1/2}} \tag{4.11a}$$

$$\mathrm{T}_0 = \int_0^\infty R(t)\mathrm{d}t = \frac{1}{4}\sum_i^m \frac{P_i}{F_i} \tag{4.11b}$$

From these equations, it is clarified that both the scales are strongly affected by the smaller frequency fluctuation, that is, by the large-scale eddy.

When it is assumed that the double correlation $R(t)$ is equal to the Lagrangean double correlation $R_\mathrm{L}(t)$, the variance that shows the degree of turbulent diffusion is expressed as

$$\begin{aligned}\sigma^2 &= 2u^2\int_0^\infty \sum_i^m \frac{(\tau - t)P_i}{1+(2\pi F_i t)^2} \\ &= u^2\sum_i^m \frac{P_i}{\gamma_i}\{-\log(\gamma_i\tau^2+1)+2(\gamma_i\tau)^{1/2}\arctan(\gamma_i\tau)^{1/2}\}\end{aligned} \tag{4.12}$$

where $\gamma_i = (2\pi)^2F_i^2$. From this equation, it is clarified that the turbulent diffusion is affected more by the larger scale eddy than the smaller scale eddy. Additionally, this variance takes the following values for the cases of $\tau \to 0$ and $\tau \to \infty$:

$$\tau \to 0: \quad \sigma^2 \cong u^2\tau^2 \tag{4.13a}$$

$$\tau \to \infty: \quad \sigma^2 \cong u^2\sum_i^m \frac{P_i}{\gamma_i}\{-2\log\tau + \pi(\gamma_i\tau)^{1/2}\} \cong u^2\pi\sum_i^m \frac{P_i}{\sqrt{\gamma_i}}\tau \tag{4.13b}$$

This relationship between the degree of diffusion by the turbulent velocity fluctuation and time is identical to the traditional general knowledge.

Let us deal with Lagrangean method and Euler method.

There are two methods to investigate the movement of fluid:

(1) *Lagrangean method*: By letting fluid particle and fluid be the mass particle and system of mass particle, respectively, the movement of each particle is expressed by time (t) and spatial coordinates (e.g., x, y, z).

(2) *Eulerian method*: The method investigates the features of fluid (e.g., velocity, pressure, density) at an arbitrary time (t) and point (x, y, z). x, y, z, and t are independent variables and velocity, pressure, density, and so on are dependent variables.

4.5 Scale-up

In general, it is consider that the scale-up of equipment is established perfectly when the distribution of ESD of the practical equipment and that of the model equipment are identical; this is because the same structure of flow is necessary to create the same phenomena in both the equipments. In Section 4.3, it is clarified that the turbulent flow field consists of a basic eddy group and its sub-harmonic eddy groups that are generated sequentially from the basic eddy group in accordance with the rule of 1/3 times wave number, that is, 3 times the size, and twice the turbulent kinetic energy. However, it is impossible that the sub-harmonic eddy group occurs infinitely, and the upper limit of the size of the eddy group is determined mainly by the size of the equipment[18]. Therefore, there should be a clear relationship between the size of the equipment and the number of eddy groups. For example, if the size of the practical equipment is set to three times that of the model equipment, a new sub-harmonic eddy group should occur; further, the turbulent flow structures of the practical and model equipment are not identical. This shows that the limit of the scale-up ratio is 3 in one dimension, that is, 27 times for the entire volume under the condition that ESD is maintained constant.

Challenge 4.2. **ESD** *of stirred vessel*

1. Scope

The ESD data in a stirred vessel has hardly been reported, and therefore, it must be clarified. Further, whether the new ESD defined by Eq. (4.7) can apply to the measured results must also be clarified.

2. Aim

To clarify the ESD distributions in a stirred vessel and ascertain whether the new ESD defined by Eq. (4.7) can be applied to the measured results.

3. Experiments

(a) Apparatus

Stirred vessel: Figure 4.4 (Cylindrical flat bottom, $D_T = 6,\ 18,\ 54$ cm).
Impeller: FBDT impeller.

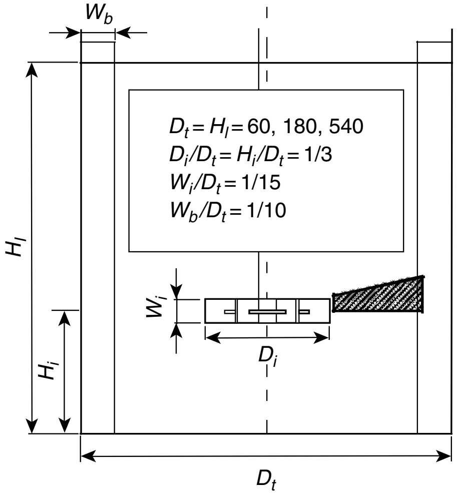


Figure 4.4 Velocity-measured region in a stirred vessel.

(b) Condition

Fluid: ion exchange water and 10 wt% aqueous glycerin solution (involves 0.5 mole/l KCl, 3×10^{-3} mole/l $K_4Fe(CN)_6$ and $K_3Fe(CN)_6$).

Re: 10 000.

Object: impeller discharge flow region.

(c) Velocity measurement method

Probe: electrode reaction velocity meter.

(d) Procedure

After it is confirmed that the flow in the vessel has attained a steady state under a fixed *Re*, the velocity fluctuations are measured at the fixed impeller discharge flow region by using an electrode reaction velocity meter. The measured ESD is fitted by the new ESD defined by Eq. (4.7).

4. Experimental results

Figure 4.5 (Measured ESD and theoretical ESD).

5. Noteworthy point

(a) The measured ESD is sufficiently expressed by the new ESD defined by Eq. (4.7).
(b) A new eddy group appears at intervals of three times the tank diameter, and the estimation of the scale-up described above is certified to be correct.

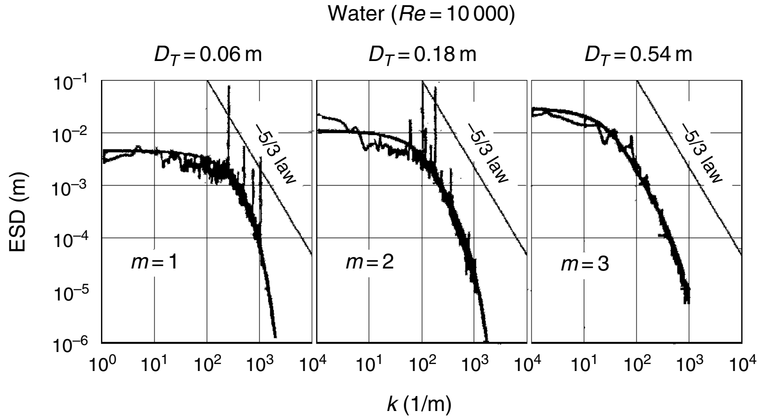


Figure 4.5 ESD in impeller discharge flow region in a stirred vessel.

(c) In the case of a stirred vessel too, the limit of the scale-up ratio is 3 in one dimension, that is, 27 times in volume under the condition that ESD is maintained constant.

6. Supplementary point

(a) Almost the same result regarding the relationship between the vessel diameter and the number of eddy groups is obtained when aqueous glycerin solution is used as the test fluid. The relationship between the average wave number of the smallest eddy group K_1 and the kinetic viscosity ν is obtained as shown in Figure 4.6.

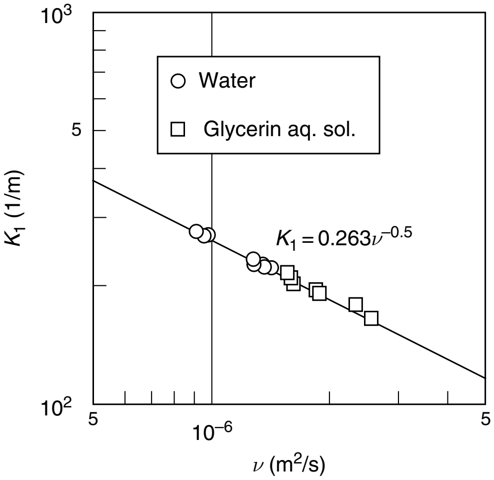


Figure 4.6 Relationship between average wave number of smallest eddy group and kinetic viscosity.

Challenge 4.3. Reliability of traditional scale-up rule of a stirred vessel [11–13,18,19]

1. Scope

Many scale-up rules have been used in industries. However, the reliability of these rules has not been discussed. In order to scale up equipment without causing any problem, it is indispensable to verify the reliability of these rules. Additionally, in this section, a new scale-up rule is proposed based on the viewpoint of information entropy.

2. Aim

To discuss the reliability of the traditional scale-up rules from the viewpoint of the new ESD defined by Eq. (4.7) and to present a new scale-up rule based on the viewpoint of information entropy.

3. Calculation

(a) Condition

Traditional scale-up rules: Table 4.2 (These scale-up rules have been used for many types of equipment scale-ups, e.g., the stirred vessel for mixing of the fluids that have no reaction ability. It is clarified that the distribution of the dimensionless turbulent statistical values by making use of impeller tip velocity $U_T(= ND)$ becomes identical regardless of impeller rotational

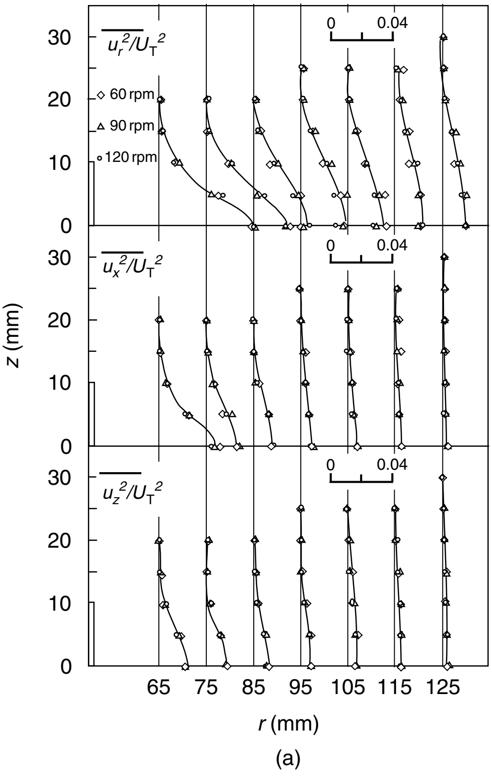
Table 4.2 Traditional scale-up rules.

ND^X = Const. Value of X	u^2D^Y = const. Value of Y	Rules	Processes
0	−2	Const. impeller revolutional speed Const. circulation time Const. impeller discharge flow rate per unit vessel volume	Fast reaction
2/3	−2/3	Const. (power) dissipation energy per unit vessel volume Const. impeller discharge flow energy	Turbulent dispersion Gas–liquid operation Reaction requiring microscale mixing
1	0	Const. impeller tip speed Const. torque per unit vessel volume	
2	2	Const. Reynolds number Const. impeller discharge flow momentum Const. torque per unit discharge flow rate	

speed N, as shown in Figure 4.7(a) and (b) (these result are obtained by measuring the velocity fluctuations at the impeller discharge flow region by using an electrode reaction velocity meter of 0.6 mm ϕ, a platinum electrode probe, and a three-dimensional velocity meter of 10 mm ϕ); Therefore, the impeller rotational speed N in the traditional scale-up rules can be replaced by using u^2 and D as $(u^2)^{1/2}/D$. In Figure 4.8, ESD values are shown based on the case of $m = 1$ by changing the value of u^2 for the respective scale-up rule. (The curve of $m = i$ is the result that is obtained by setting u^2 as 3^{i-1} times that in the case of $m = 1$.))

(b) Method

According to the scale-up rule, the kinetic energy and the ESD distribution corresponding to this value of the kinetic energy is calculated under the assumption that a new eddy group appears every three times the vessel diameter.



(a)

Figure 4.7 (a) Distributions of energy values of turbulent fluctuations (b) distributions of double correlation values of turbulent fluctuations in impeller discharge flow region in a stirred vessel.

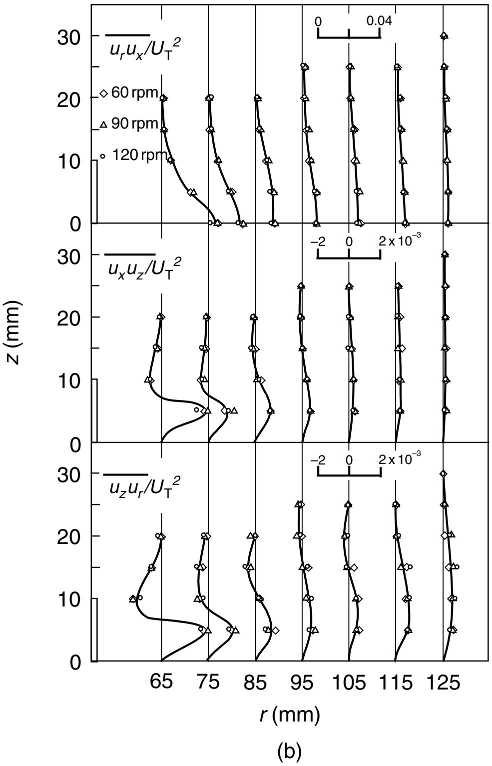


Figure 4.7 (Continued)

The reliability of the scale-up rule is evaluated based on the relationship between the number of eddy groups and ESD. If there is an overlapping wave number region, regardless of the number of eddy groups, the scale-up rule has a reliability when the overlapping wave number range is significant for the observed phenomena. (It is also true that when the scale-up ratio is less than 3, it is possible to scale up by maintaining a constant value of u^2, and there is no change in the ESD value before and after the scale up.)

4. Calculated results

Figure 4.8 (EK_1/u_1^2 versus dimensionless wavenumber).

5. Noteworthy point

(a) $ND^0 = constant$ has no reliability.

This rule is equivalent to u^2D^{-2} = constant. Since there is no overlapping in any wave number range, there is no reliability.

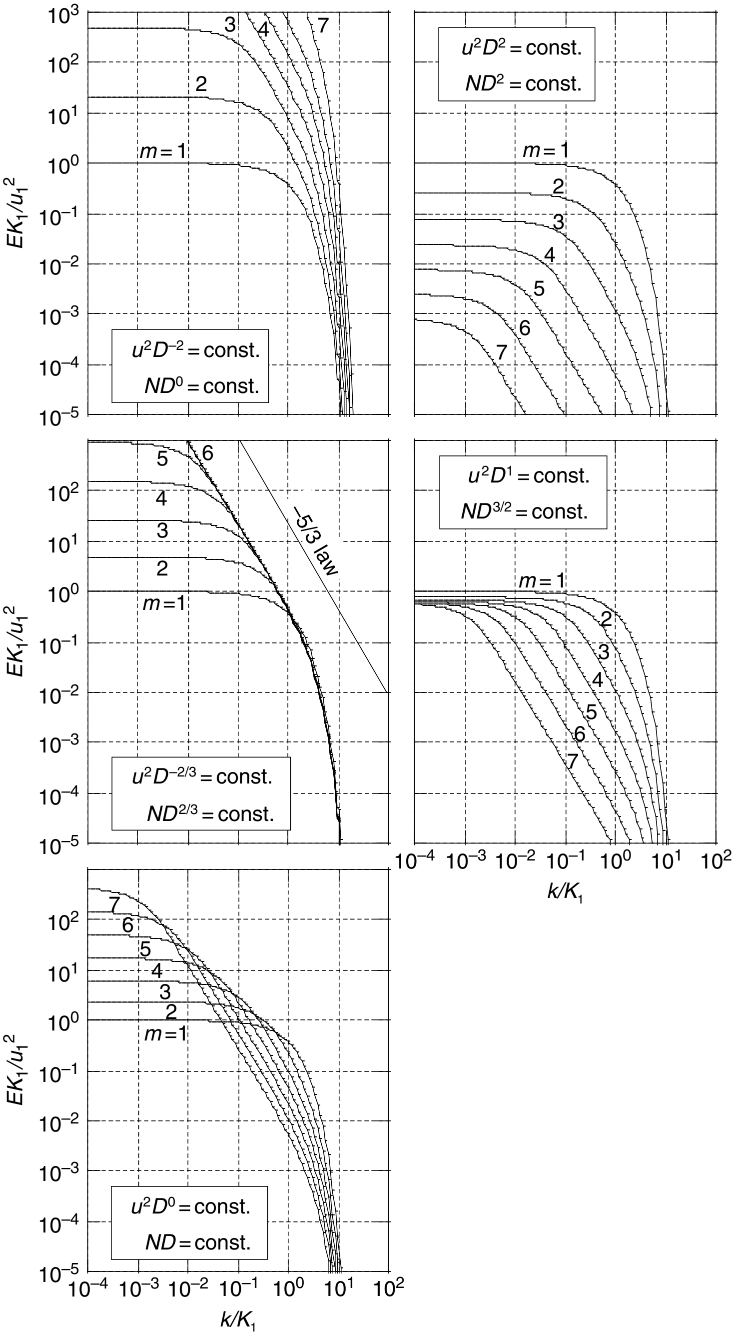


Figure 4.8 Evaluation of traditional scale-up rules based on new ESD function.

(b) $ND^{2/3} = constant$ has reliability under a certain condition.

This is equivalent to $u^2D^{-2/3}$ = constant. All the curves overlap each other in the higher wave number range in which Kolmogorov's −5/3 law can be applied; however, there is no overlapping in the lower wave number range. This shows that this rule is reliable when the wave number range in which Kolmogorov's −5/3 law can be applicable is significant for the observed phenomena.

(c) $ND = constant$ has no reliability.

This scale-up rule is equivalent to u^2D^0 = constant. Since all the curves intersect at only one point, the reliability of this scale-up rule is poor.

(d) $ND^2 = constant$ has no reliability.

This scale-up rule is equivalent to u^2D^2 = constant. Since there is no overlapping in any wave number range, it can be stated that there is no reliability for this scale-up.

(e) $ND^{3/2} = constant$ has reliability under a limited condition.

This scale-up rule is not listed in Table 4.2, and it is equivalent to u^2D^1 = constant. However, there is no overlapping in the higher wave number range, while almost all the curves overlap in the lower wave number range. This shows that this scale-up rule has reliability when the lower wave number range is significant for the observed phenomena.

Challenge 4.4. Scale-up of circular pipe

1. Scope

The pipe has not been treated as a chemical equipment but only as a means for transportation. However, recently, line mixing and inline reaction have been considered, and the scale-up rule for pipes comes into question. Therefore, the relationship between the pipe diameter and turbulent flow structure in a circular pipe must be clarified.

2. Aim

To clarify the relationship between the pipe diameter D and turbulent flow structure, in other words, the number of eddy groups m.

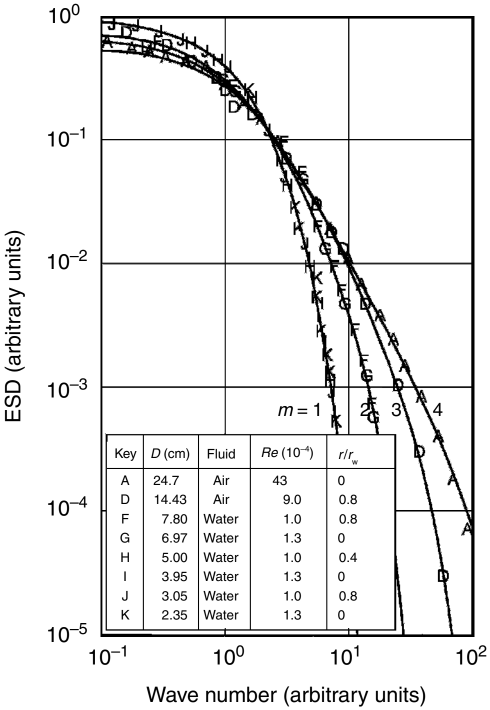
3. Calculation

(a) Condition

Data quoted: Figure 4.9 (Air flow and water flow in a circular pipe, involve the data by authors).

(b) Method

Under the assumption that a new eddy group appears with three times the vessel diameter, the ESD that fits the measured ESD data is determined



Key	D (cm)	Fluid	Re (10^{-4})	r/r_w
A	24.7	Air	43	0
D	14.43	Air	9.0	0.8
F	7.80	Water	1.0	0.8
G	6.97	Water	1.3	0
H	5.00	Water	1.0	0.4
I	3.95	Water	1.3	0
J	3.05	Water	1.0	0.8
K	2.35	Water	1.3	0

Figure 4.9 ESD for air and water flow in a circular pipe (involve the data by authors).

by changing the value of m. Based on the relationship between D and m, the average wave number of the basic eddy group of the respective fluids is estimated.

4. Calculated results

Figure 4.10 (D versus m).

$K_{1W} = 1.40\,\text{cm}^{-1}$
$K_{1A} = 5.70\,\text{cm}^{-1}$

Table 4.3 (The region of pipe inner diameter in which the value of number of eddy groups does not change).

5. Noteworthy point

(a) m depends on D and the fluid. By considering that m increases by one when D becomes 3 times, the relationship between D and m for the respective fluid is shown in Figure 4.10.

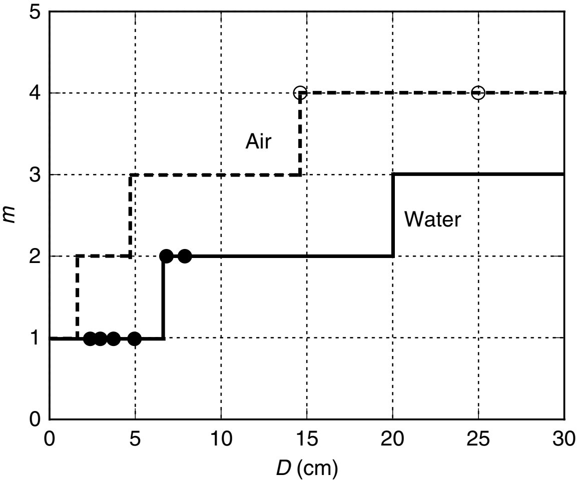


Figure 4.10 Relationship between pipe diameter (cm) and number of eddy groups.

Table 4.3 Relationship between pipe inner diameter (cm) and number of eddy groups.

m	Water	Air
1	$0 < D \leq 6.6$	$0 < D \leq 1.62$
2	$6.6 < D \leq 19.8$	$1.62 < D \leq 4.86$
3	$19.8 < D \leq 59.4$	$4.86 < D \leq 14.6$
4	$59.4 < D \leq 178$	$14.6 < D \leq 43.8$
5	$178 < D \leq 535$	$43.8 < D \leq 131$
6	$535 < D \leq 1604$	$131 < D \leq 394$

(b) The average wave number of the basic eddy group for the used air and water can be calculated based on Figure 4.8 as

$$K_{1\mathrm{W}} = 1.40\,\mathrm{cm}^{-1}, \quad K_{1\mathrm{A}} = 5.70\,\mathrm{cm}^{-1}$$

(c) Based on the value described above, it is possible to obtain the space scale, that is, D that corresponds to the average wave number. The relationship between D and m can be derived from the average space scale of the basic eddy group, as shown in Table 4.3. Indeed, this table shows the standard of scale-up of the circular pipe for air and water.

4.6 Energy spectrum density distribution function of non-Newtonian liquid

Until now, we have focused on Newtonian fluids. However, there are many kinds of non-Newtonian fluids in chemical industries, and the turbulent flow structure of non-Newtonian fluids has remained unknown. It is very important to solve this problem. In this section, the power law fluid is considered to discuss the turbulent flow structure because the power law fluid is a representative non-Newtonian fluid and it is easy to treat it mathematically.

The rheological equation of the state of the power law fluid is expressed as

$$\tau = A\gamma^n \tag{4.14}$$

where τ is the shear stress and γ the shear rate. The turbulent structure of the power law fluid must be compared with that of the Newtonian fluid. However, it is assumed that there is no difference in the relation among the forces with regard to turbulent eddies in the power law fluid and Newtonian fluid. In other words, the momentum that is accepted by the eddy due to a unit area of surface is proportional to the shear stress that occurs in a unit surface area as

$$\rho u^2 \propto \tau \tag{4.15}$$

The relationship shown in the above equation can be rewritten as follows for the respective fluids:

$$\text{Newtonian fluid} \quad \rho_\mathrm{N} {u_\mathrm{N}}^2 \propto \mu_\mathrm{N} \frac{\mathrm{d}u_\mathrm{N}}{\mathrm{d}r} \propto \mu_\mathrm{N} \left(\frac{u_\mathrm{N}}{\ell}\right)^2 \tag{4.16}$$

$$\text{Power law fluid} \quad \rho_\mathrm{P} {u_\mathrm{P}}^2 \propto \mu_\mathrm{P} \frac{\mathrm{d}u_\mathrm{P}}{\mathrm{d}r} \propto \mu_\mathrm{P} \left(\frac{u_\mathrm{P}}{\ell}\right)^2 \tag{4.17}$$

The density ρ and viscosity μ are physical properties whose values are fixed. Additionally, the eddy diameter ℓ can be considered to be proportional to $1/k$. Therefore, Eqs (4.16) and (4.17) can be rewritten as

$$\text{Newtonian fluid} \quad {u_\mathrm{N}}^2 \propto k^2 \tag{4.18}$$

$$\text{Power law fluid} \quad {u_\mathrm{P}}^2 \propto k^{2n/(2-n)} \tag{4.19}$$

On the other hand, it is clarified that the following relation exists between the kinetic energy of a Newtonian fluid and that of a power law fluid:

$$\frac{{u_\mathrm{P}}^2}{{u_\mathrm{N}}^2} \propto k^{[4(n-1)]/(2-n)} \tag{4.20}$$

By considering that u^2 here corresponds to $u_N{}^2$ in the above discussion, the following equation can be obtained by substituting Eq. (4.20) in Eq. (4.19).

$$\frac{E(k)}{u_P{}^2} = \frac{1}{K_1 \sum_1^m 2^{j-1}} \sum_1^m \left\{ 6^{j-1} \exp\left(-3^{j-1} \frac{k}{K_1} \right) \right\} (Bk^{[4(n-1)]/(2-n)}) \tag{4.21}$$

where B is a coefficient that standardizes the expression as an ESD function.

Equation (4.21) is the only ESD function for the power law fluid.

Further, it is also possible to derive the same result by considering the energy. In other words, when the energy that is accepted by the eddy through a unit surface area is proportional to the energy that is dissipated over a unit eddy surface area, the following relation can be obtained:

$$u\frac{1}{2}\rho u^2 \propto u\tau \tag{4.22}$$

This equation corresponds to Eq. (4.15).

Challenge 4.5. ESD of power law fluid

1. Scope

However, many kinds of non-Newtonian fluid exist in chemical industries, and the turbulent flow structure of a non-Newtonian fluid is yet to be understood. ESD data are few, and knowledge about ESD is insufficient. It is very important to at least clarify ESD of non-Newtonian fluids.

2. Aim

To clarify whether the new ESD defined by Eq. (4.21) can be applied for a power law fluid.

3. Experiments

(a) Apparatus

Stirred vessel: Figure 4.4 (Cylindrical flat bottom vessel, four baffles).
Impeller: Figure 2.4 (FBDT impeller).
Fluid: 0.2–0.9 wt% aqueous CMC solution (involves 0.5 mole/l KCl, 3×10^{-3} mole/l $K_4Fe(CN)_6$ and $K_3Fe(CN)_6$, $\rho_P = 1.0\,\text{g/cm}^3, n = 0.817\text{–}0.999,\ A = 0.0029\text{–}0.1032,\ \tau = A\gamma^n = 0.0283\gamma^{0.891}$ in the case of 0.6 wt% (Figure 4.11)).

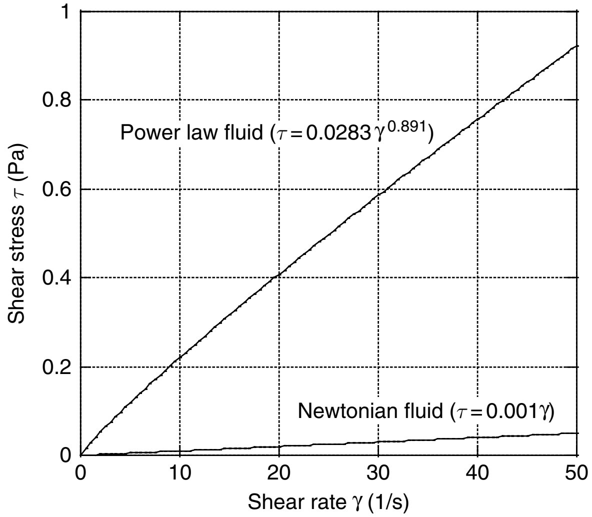


Figure 4.11 Rheology characteristics of 0.6 wt% aq. CMC sol.

Re: 5000.
Object: impeller discharge flow region.

(b) Velocity measurement method

Probe: electrode reaction velocity meter.

(c) Procedure

After it is confirmed that the flow in the vessel attains a steady state under a fixed *Re*, the velocity fluctuations are measured at the fixed impeller discharge flow region by using an electrode reaction velocity meter. The measured ESD is fitted by the new ESD defined by Eq. (4.21).

4. Experimental results

Figure 4.12 (Measured result of 0.6 wt% aqueous CMC solution and fitted distribution of ESD).

5. Noteworthy point

(a) The new ESD defined by Eq. (4.21) can sufficiently fit the measured ESD.

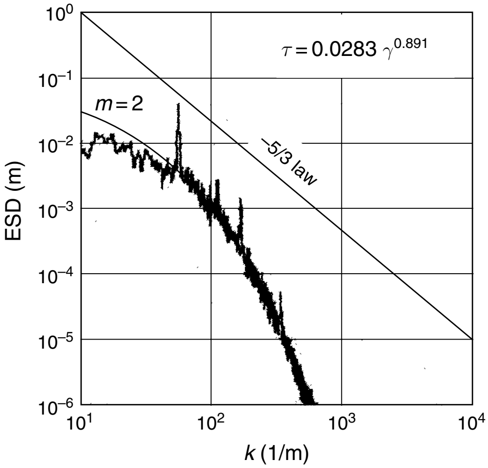


Figure 4.12 Measured ESD of 0.6 wt aq. CMC sol. and fitted ESD curve based on new ESD function.

6. Supplementary point

(a) In the case of other concentrations of CMC, the tendency of the results is almost identical.

Let us deal with the electrode reaction method controled by mass transport rate and its probes.

The test liquid is an aqueous solution of 3×10^{-3} mole/l of $K_4Fe(CN)_6$ and $K_3Fe(CN)_6$ and 5×10^{-1} mole/l of KCl. In this case, the ratio of the boundary layer thickness of momentum and concentration is about 10 to 1. When the main velocity changes, the thickness of the momentum boundary layer first changes and then, the thickness of the concentration boundary layer changes. Finally, the diffusion current changes according to the change in the main flow. The oxidation of the ferrocyanide ion is used as the electrolyte reaction on the platinum electrodes. The liquid velocity can be measured with three types of probes, as shown in Figure 4.13. Both the platinum sphere and cylindrical probes are used for measuring the absolute value of the velocity vector continuously, and the multi-electrode probe is used for measuring the three-dimensional instantaneous velocity vector continuously. In the case of the multi-electrode probe, the distance between the stagnation point and the electrode depends on the arrangement of the electrode, that is, the thickness of the boundary layer is different, and it is possible to determine the direction of the velocity vector.

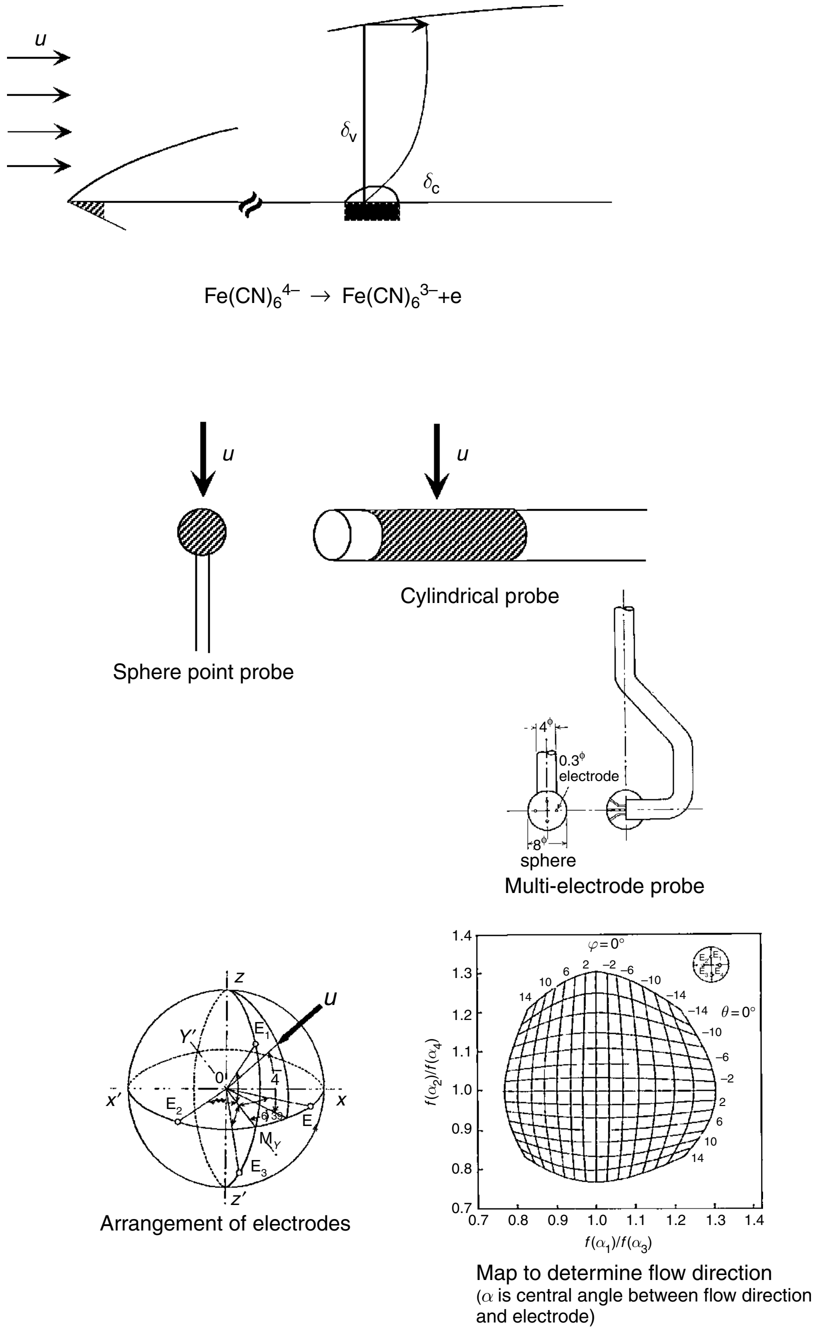


Figure 4.13 Velocity-measuring probes based on electrode reaction controlled by mass transfer rate.

4.7 Summary

In this chapter, the following points have been clarified:

1. By using the information entropy, a new expression of the ESD function for a wide wave number range is presented.
2. The turbulent flow field consists of the basic eddy group and its sub-harmonic eddy groups that are generated sequentially from the basic eddy group in accordance with the rule that the wave number and the kinetic energy of $i+1$-eddy group are, respectively, 1/3 times and 2 times those of the i-eddy group.
3. The limit of the scale-up ratio is 3 in one dimension, that is, 27 times in volume under the condition that ESD is maintained constant. However, in the case of a stirred vessel, the rule $ND^{3/2} = \text{constant}$ has reliability under a limited condition. (Since almost all the curves overlap at the lower wave number range, this fact shows that this scale-up rule has reliability when the lower wave number range is significant for the observed phenomena.)
4. A new scale-up rule for a circular pipe is presented.
5. ESD for the power law fluid is presented.
6. The usefulness of the newly defined ESD is confirmed by applying it to several examples and subjects as follows:
    - (a) Measured ESD and fitting curves based on newly defined ESD function,
    - (b) ESD of stirred vessel,
    - (c) reliability of traditional scale-up rule of a stirred vessel,
    - (d) scale-up of circular pipe,
    - (e) ESD of power law fluid.

This page intentionally left blank

CHAPTER 5

Particle Size Distribution

5.1 Introduction

Particle treatment can be observed many times in chemical industries. In mixing operations, the particles of the disperse phase are very important. The first aim in mixing by using impellers in liquid–liquid mixing is to disperse the fine droplets in a vessel. Knowing the size distribution of droplets is indispensable to establish a reasonable design of equipment because the state of dispersion has a close relationship with the operation conditions. In order to clarify the mass transfer in a two-phase contactor, such as a vessel for liquid–liquid mixing, it is indispensable to obtain a clear particle size distribution of the dispersion phase. There are some distribution functions for expressing practical particle size distributions. The most widely utilized distribution function is the Rosin–Rammler probability density distribution function. Crushed product and powder dust are expressed by Rosin–Rammler probability density distribution function because the particle size has a wide range. This distribution function usually applies to data that are too skewed to be fitted as a log-normal size probability density distribution function. It has been said that the crystal size distribution can be expressed sufficiently by this log-normal probability density distribution function. Further, in addition to the Rosin–Rammler probability density distribution function, a log-normal probability density distribution function and normal probability density distribution function as a general particle size distribution function exist. Droplets produced by liquid–liquid mixing and bubbles produced by gas–liquid mixing are often satisfactorily represented by the normal probability density distribution function. The particle size distribution in the case of gas–liquid jet mixing vessel shows a sharp normal distribution. The log-normal probability density distribution function often applies to naturally occurring powders, the product of crushing, and so on, and is the widely employed size distribution function in all practical works. The particle size distribution in the case of the gas–liquid mixing and bubble column shows log-normal probability density distribution function. Thus, even if the system is fixed, there is no uniform view about the expression of particle size probability density distribution (PSD). It is noteworthy that these traditional distribution functions are only mathematical expressions and have no physical basis and significance, which is their biggest disadvantage. Additionally, there is a possibility that the following inconvenient situation will occur. When the PSDs under different operation conditions are fitted by using different fitting functions, it is difficult to relate the fitting parameters and operation conditions. In other words, it becomes impossible to estimate the operation condition that will produce the expected PSD. Therefore, it is indispensable to build a general PSD that can sufficiently fit every PSD. However, every traditional function described above cannot satisfy this condition. Of course, the expected function should be better than the traditional equation as described above. When the process of particle production is considered, the PSD should be treated by probability terms. In this chapter, a new general expression of the PSD function is discussed by making use of information entropy.

Let us deal with the relationship between the PSD and probability distribution.

The particle size distribution can be plotted in terms of the cumulative percent oversize or undersize in relation to the particle diameters. The weight, volume, number, and so on are used for percentage. By differentiating the cumulative distribution with respect to the diameter of the particle, the PSD can be obtained.

*5.2 Particle size probability density distribution function (*PSD *function)*

The droplets in liquid–liquid mixing and bubble in gas–liquid mixing are broken when the outer force/energy such as the shear stress based on the shear rate exceeds the inner force/energy such as surface tension. In the crystallization, the crystal size depends on the crystal growth rate, which is affected by temperature, concentration, pH, and so on. However, when it is assumed that the secondary nucleus is produced from the first nucleus by collision with an impeller or shear stress growth without a change in the particle size distribution, the crystal size distribution can be determined by a balance of the outer force/energy and inner force/energy. Additionally, crushed products can be considered to be produced after crushing when the outer force/energy exceeds the inner force/energy. In the following discussion, the droplet, bubble, crystal, and crushed product size probability density distribution function (hereafter PSD) are discussed based on information entropy. Henceforth, droplet, bubble, crystal, and crushed product[1–3] are called particles in the lump.

(1) Original variable in discussion of PSD

The following balance equation is established for the critical particle in all the cases described above:

$$F_O S = F_I \ell \quad \text{or} \quad F_O V = F_I S \tag{5.1}$$

where F_O is the outer force such as shear stress: F_I the inner force such as surface tension; S the surface area of a particle; ℓ the diameter of a particle; and V the volume of a particle. The internal force/energy does not change by changing the operation conditions such as the impeller rotational speed. The difference in operation condition has effect only on the values of external force/energy. Therefore, in general, ℓ/S or S/V (specific surface) will be a suitable variable to examine PSD:

$$\frac{\ell}{S} = \frac{F_O}{F_I} \quad \text{or} \quad \frac{S}{V} = \frac{F_O}{F_1} \tag{5.2}$$

Both ℓ/S and S/V are proportional to $1/\ell$ and Eq. (5.2) is rewritten as

$$\frac{1}{\ell} \propto \frac{F_O}{F_I} \tag{5.3}$$

From this relational equation, it can be said that the variable for discussing PSD should be $1/\ell$.

(2) Definition of PSD from the viewpoint of information entropy

Next, the definition of PSD is discussed based on the uncertainty regarding "the size of the particle that is selected" when a particle is selected. The original *PSD* $q_O{}^*(1/\ell)$ is regarded as a probability density function, and the PSD satisfies the following standardization condition:

$$\int_0^\infty q_O{}^*(1/\ell)\mathrm{d}(1/\ell) = 1 \tag{5.4}$$

The amount of uncertainty regarding "the size of the particle that is selected" is expressed by the information entropy as

$$H(1/\ell) = -\int_0^\infty q_O{}^*(1/\ell)\log q_O{}^*(1/\ell)\mathrm{d}(1/\ell) \tag{5.5}$$

In order to define PSD from the viewpoint of information entropy, the following two assumptions are made to discuss the PSD expression:

(1) The PSD function $q_O{}^*(1/\ell)$ yields the maximum amount of information entropy.
(2) There is an average value of $1/L$:

$$\int_0^\infty q_O{}^*(1/\ell)\mathrm{d}(1/\ell) = 1/L. \tag{5.6}$$

Based on these assumptions, the original PSD is obtained as follows according to Section 1.7.

$$q_O{}^*(1/\ell) \approx \frac{1}{1/L}\exp\left(-\frac{1/\ell}{1/L}\right) \tag{5.7}$$

Since the variable $1/\ell$ is not commonly used in chemical engineering, the most common variable ℓ is used. When the variable is ℓ, Eq. (5.7) becomes

$$q_{\mathrm{O}}(\ell) \approx L \exp\left(-\frac{L}{\ell}\right) \tag{5.8}$$

In fact, this $q_{\mathrm{O}}(\ell)$ corresponds to the original PSD of the droplets, bubbles, crystals, crushed products, and so on. The limiting values of Eq. (5.8) are expressed as

$$\lim_{\ell \to 0} q_{\mathrm{O}}(\ell) = 0$$
$$\lim_{\ell \to \infty} q_{\mathrm{O}}(\ell) = L$$

However, in the original PSD derived theoretically as described above, it is difficult to expect that every particle of each size actually appears with a probability of 100%. In the following discussion, the probability of the appearance of the particle is termed as the realizable probability. The large particle can be easily broken up by an external force/energy. On the other hand, it is difficult for the small particle to receive the effect of an external force/energy since it can be easily hidden in the main stream in a stirred vessel. This implies that a larger particle is unlikely to exist, while a smaller particle is likely to exist. In other words, the realizable probability takes a smaller value as the particle size increases. When the realizable probability is considered to depend on the factor Q that represents the characteristics of the surface area of a particle, the realizable probability can be set as a function of Q as $P(Q)$. By considering that the value of the gradient of $P(Q)$ with respect to the factor Q has a negative value, the realizable probability $P(Q)$ has a relationship with the probability density function $p(Q)$ as

$$p(Q) = -\frac{\mathrm{d}P(Q)}{\mathrm{d}Q} \tag{5.9}$$

The probability density function and $p(Q)$ obviously satisfy the following standardization condition as

$$\int_0^\infty p(Q)\mathrm{d}Q = 1 \tag{5.10}$$

As follows, when a particle is selected, PSD is introduced based on the uncertainty regarding "the value of Q of the particle that is selected." The amount of uncertainty is expressed by the information entropy as

$$H = -\int_0^\infty p(Q) \log p(Q)\mathrm{d}Q \tag{5.11}$$

In order to determine the expression of $p(Q)$, the following two assumptions are made:

(1) $p(Q)$ yields the maximum amount of information entropy.
(2) There is an average value of Q_A:

$$\int_0^{\infty} Qp(Q)\mathrm{d}Q = Q_A. \tag{5.12}$$

Based on these assumptions, $p(Q)$ is obtained as follows, according to Section 1.7:

$$p(Q) = \frac{1}{Q_A}\exp\left(-\frac{Q}{Q_A}\right) \tag{5.13}$$

By substituting Eq. (5.13) in Eq. (5.9), the realizable probability $P(Q)$ can be obtained as

$$P(Q) = 1 - \int_0^{Q} p(Q)\mathrm{d}Q = \exp\left(-\frac{Q}{Q_A}\right) \tag{5.14}$$

When $1/Q_A$ is replaced by B, the above equation can be rewritten as

$$P = \exp(-BQ) \tag{5.15}$$

It is now assumed that the factor Q has the following relation with the particle size ℓ:

$$Q \propto \left(\frac{\ell}{L}\right)^{C} \tag{5.16}$$

Based on Eqs (5.15) and (5.16), the realizable probability $P(Q)$ can be deduced as

$$P(Q) \approx \exp\left\{-B\left(\frac{\ell}{L}\right)^{C}\right\} \tag{5.17}$$

Finally, the actual PSD is shown by the following equation that is obtained as the product of the original PSD shown in Eq. (5.7) and the realizable probability shown in Eq. (5.17):

$$q(\ell) = q_O(1/\ell)P(Q) = AL\exp\left(-\frac{L}{\ell}\right)\exp\left\{-B\left(\frac{\ell}{L}\right)^{C}\right\} \tag{5.18}$$

where A is the coefficient that satisfies the following standardized condition for PSD:

$$\int_0^{\infty} q\mathrm{d}\ell = 1 \tag{5.19}$$

When C has a value of 2, the factor Q corresponds to the surface area of a particle.

(3) PSD based on ESD

Droplets, bubbles, and crystals are formed together in a stirred vessel. Therefore, it is possible to consider the energy derived from ESD as the external energy that contributes to the break-up of particles. An eddy larger than the observed particle swallows the particle, and the particle circulates only in the vessel. On the other hand, if an eddy is extremely small when compared to the observed particle, it is difficult to induce an effect on the particle for causing break-up. In other words, an eddy particle with a size that is almost the same as that of the particle will primarily affect the particle break-up. Therefore, the energy of the eddy size, which is the same as that of the particle, is considered as the external energy. The external energy F_{O}, which has the dimension of [kg/(m s^2)], for example, is expressed by the following equation:

$$F_{\mathrm{O}} \approx \rho E_k \Delta k \tag{5.20}$$

where E_k is the turbulent energy of the eddy that has the same size as the observed particle and ρ is the density of the continuous phase, for example, the liquid. On the other hand, the internal energy F_{I}, such as surface tension, which has the dimension of [N/m] = [kg/s^2], is considered. Therefore, the following balance equation is established for the critical particle:

$$\rho E_k \Delta k V_T \approx F_{\mathrm{I}} \ell^2 N \tag{5.21}$$

where V_T is the volume of the fluid and N the number of particles in the stirred vessel.

It is possible to obtain the total volume of the particle under the condition of V_T = constant, F_{I} = constant, and ρ = constant. Next, the total volume of the particle that has the size of ℓ can derived as

$$\ell^3 N \approx \frac{\rho V_T}{F_{\mathrm{I}}} E_{k=1/\ell} \Delta k \ell = \frac{\rho V_T}{F_{\mathrm{I}}} \frac{\Delta k}{k} E_{k=1/\ell} \tag{5.22}$$

Finally, the original eddy size distribution can be obtained as

$$q_{\mathrm{O}}(\ell) \propto \ell^3 N \propto \frac{\rho V_T}{F_{\mathrm{I}}} \Delta k \ell E_{k=1/\ell} \approx \frac{\rho V_T}{F_{\mathrm{I}}} \frac{\Delta k}{k} E_{k=1/\ell} \tag{5.23}$$

Since it is possible to assume that $(\rho V_T/F_I)(\Delta k/k)$ has a constant value, the following equation can be obtained:

$$q_{\mathrm{O}}(\ell) \propto E_k \propto \sum_{i=1}^{m} L_i \exp\left(-\frac{L_i}{\ell}\right) \tag{5.24}$$

It is natural to assume that the turbulent energy does not depend on the eddy group. It is clarified that PSD based on the viewpoint described above can be written as the following equation, when the realizable probability is introduced, as was the case in the previous section:

$$q(\ell) \propto q_{\mathrm{O}}(\ell)P(Q) \propto \sum_{i=1}^{m} L_i \exp\left(-\frac{L_i}{\ell}\right) \exp\left\{-B\left(\frac{\ell}{L}\right)^{C}\right\} \tag{5.25}$$

The style of this equation is isomorphic with that of the aforesaid PSD shown in Eq. (5.18) without a difference in the existence of sigma. However, the difference is not as significant from a practical viewpoint, where the distribution for the case $i = 1$ is considered to have the highest weight.

5.3 Eddy size distribution in a turbulent flow

It is possible to apply the same consideration to the eddy size distribution in a turbulent flow because eddies in a turbulent flow are produced by the impeller, shear stress, and so on, without artificiality. There is no difficulty in considering the following relationships between the wavenumber k and the diameter of eddy ℓ in a turbulent flow:

$$L \propto 1/k \qquad L_i \propto 1/K_i \tag{5.26}$$

By using these relationships, ESD shown in Eq. (4.7) can be rewritten as

$$\frac{E(k)}{u^2} = \sum_{i=1}^{m} \frac{E_i(k)}{u_i^2} = \sum_{i=1}^{m} \frac{1}{u_i^2}\frac{1}{K_i} \exp\left(-\frac{k}{K}\right) = \sum_{i=1}^{m} \frac{1}{u_i^2} L_i \exp\left(-\frac{L_i}{\ell}\right) \tag{5.27}$$

The number and volume of the respective scale of eddies can be obtained by using Eq. (5.27), and the eddy size distribution can be shown in the following equation by assuming that the energy per unit volume of the eddy is equal irrespective of the eddies:

$$q_{\mathrm{e}}(\ell) \approx \sum_{i=1}^{m} \frac{E_i}{u_i^2 \ell^3} \ell^3 = \sum_{i=1}^{m} \frac{1}{K_i} \exp\left(-\frac{k}{K_i}\right) \propto \sum_{i=1}^{m} L_i \exp\left(-\frac{L_i}{\ell}\right) \tag{5.28}$$

The equation style is isomorphic with that of the aforesaid original particle PSD without a difference in the existence of sigma. This fact introduces the following understanding. An eddy is segmented when it strikes impellers and baffles or is subjected to turbulent shear forces. In other words, it is possible to deduce ESF from the same viewpoint for each eddy group as that in the aforesaid approach with regard to the original PSD. However, in the case of droplets, bubbles, crystals, and crushed products, there is no particle that involves another smaller particle except in rare cases such as emulsion of oil/water/oil, and it is common for an eddy to involve another smaller eddy in the case of a turbulent flow field. This is one reason why it is not necessary to consider a realizable probability in the case of a turbulent flow field.

Challenge 5.1. Usefulness of new PSD

1. Scope

There are a fair number of PSD functions to express the practical PSD. However, these functions are only mathematical expressions and have no physical basis or significance. Additionally, there is a problem when the practical PSD is expressed by using different PSD functions, for example, Rosin–Rammler PSD function, normal PSD function, and log-normal PSD function. This is because it becomes difficult to correlate the values of a parameter and operation condition. Therefore, it is indispensable to clarify whether a new PSD defined by Eq. (5.18) has an advantage over the traditional PSD.

2. Aim

To verify whether the new PSD defined by Eq. (5.18) can replace the traditional PSD.

3. Calculation

(a) Condition

Traditional PSD: Table 5.1 (Rosin–Rammler PSD function, normal PSD function, log-normal PSD function).

Table 5.1 Values of curve fitting parameters in new PSD function for typical three traditional PSDs.

Distribution	Generation Parameter		A	L	B	C
Rosin–Rammler	N	D				
	2	5×10^{-4}	2.83×10^{-3}	15.8	8.75×10^{-2}	2
	2.6	9.5×10^{-6}	9.87×10^{-4}	80.9	9.70×10^{-1}	2
	2.8	8×10^{-6}	2.10×10^{-3}	79.3	1.78×10^{0}	2
Normal	x_m	σ_N				
	50	10	1.51×10^{0}	305	1.24×10^{2}	2
	50	15	2.50×10^{-2}	159	1.87×10^{1}	2
	50	20	3.40×10^{-3}	80.1	2.66×10^{0}	2
	30	5	2.081×10^{2}	265	3.66×10^{2}	2
Log-normal	x_m	σ_{LN}				
"wide"	20	2	1.06×10^{-2}	6.63	5.65×10^{-2}	2
	45	1.6	2.84×10^{-3}	48.4	1.03×10^{0}	2
	70	1.5	1.96×10^{-3}	110	2.85×10^{0}	2
"narrow"	25	1.3	1.48×10^{-1}	84.5	2.30×10^{1}	2

The usefulness of the newly presented PSD shown in Eq. (5.17) is examined by the Rosin–Rammler PSD function:

$$q_{\mathrm{RR}}(x) = ndx^{n-1}\exp(-dx^n) \tag{5.29}$$

The normal PSD function is

$$q_{\mathrm{N}}(x) = \frac{1}{(2\pi\sigma_{\mathrm{N}})^{1/2}}\exp\left\{-\frac{(x-x_{\mathrm{om}})^2}{2\sigma_{\mathrm{N}}{}^2}\right\} \tag{5.30}$$

The log-normal PSD function is

$$q_{\mathrm{LN}}(x) = \frac{1}{x\{2\pi(\ln\sigma_{\mathrm{LN}})^2\}^{1/2}}\exp\left\{-\frac{(\ln x - \ln x_{\mathrm{om}})^2}{2(\ln\sigma_{\mathrm{LN}})^2}\right\} \tag{5.31}$$

(b) Method

First, the data are generated based on these three common PSDs. Secondly, these data are fitted by a new PSD defined by Eq. (5.18).

4. Calculated result

Figure 5.1 (New PSD and data generated from traditional PSD functions). Table 5.1 (Parameters and statistical fits).

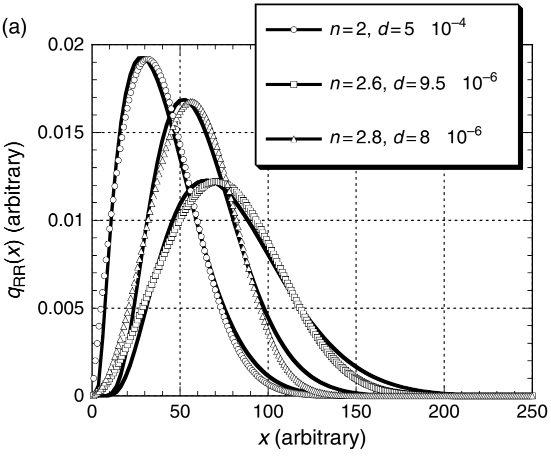


Figure 5.1 (a) Data of Rosin–Rammler distribution and fitted PSD curve based on new PSD function. (b) Original PSD curve and realized probability curve in the case of Rosin–Rammler distribution. (c) Data of log-normal distribution and fitted PSD curve based on new PSD function. (d) Data of normal distribution and fitted PSD curve based on new PSD function.

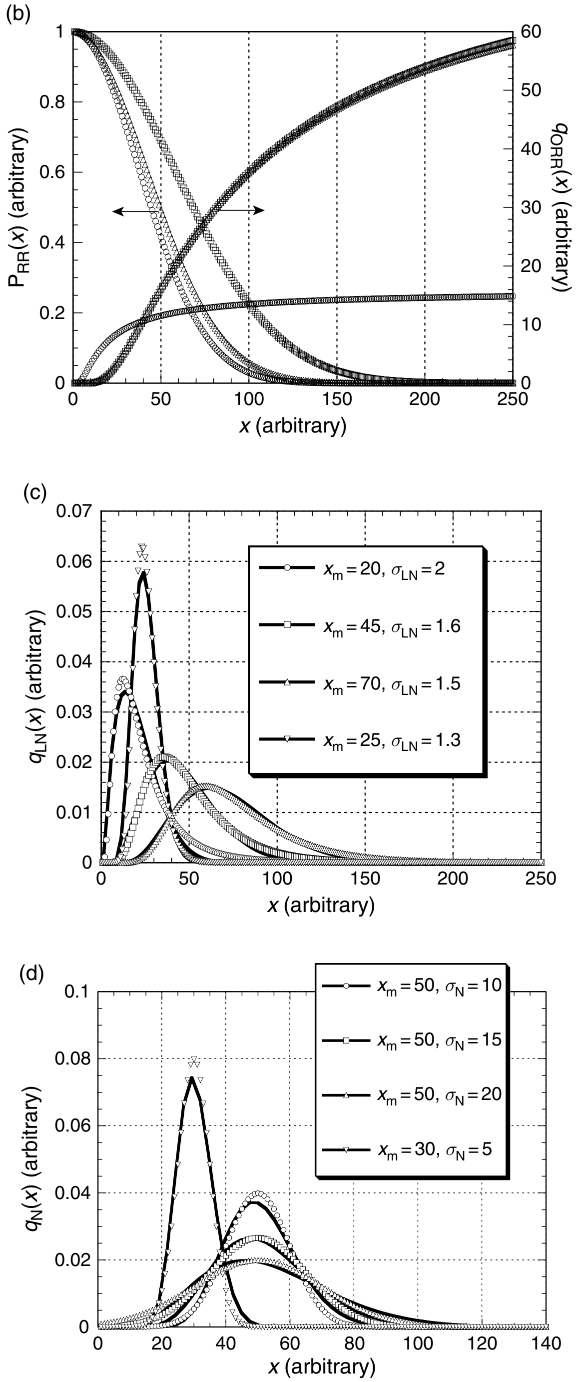


Figure 5.1 (Continued)

5. Noteworthy point

(a) The new PSD, which uses three adjustable parameters, can effectively represent both narrow and wide distributions that are described by three traditional common PSDs; in other words, the generated data and fitting curve clearly demonstrate that the fit is excellent and that the two distributions are virtually indistinguishable for all practical purposes even if the value of the exponent C is fixed as 2; this is from the consideration that the surface area greatly affects the realization of the particles. However, there is a slight difference between the data and fitting curves.

    Of course, when the value of the exponent C is another adjustable parameter, slightly better fits are obtained for every size distribution function.

(b) There is a possibility that many discussions on the actual size distribution can be simplified because the new PSD yields general common information on the original characteristics, for example, the original size distribution function and realizable probability function.

Challenge 5.2. Droplet size distribution in liquid–liquid mixing

1. Scope

Liquid–liquid mixing has been widely used in chemical industries. The state of dispersion is determined by the balance of the break-up and coalescence of droplets. In the case of liquid–liquid mixing, the break-up of the droplet is accelerated in the impeller region. Although the droplet size distribution in the operation has been expressed by various PSD functions, the PSD function that is utilized the most is the normal PSD function. However, there is no physical background to apply the normal PSD function to the droplet size distribution. Additionally, when the droplet size distribution is expressed by various PSD functions, it becomes difficult to discuss the relationship between the parameters in PSD and operation conditions. This is one of the obstacles for developing particle technology.

2. Aim

To examine whether the new PSD defined by Eq. (5.18) can sufficiently express droplet size distribution in liquid–liquid mixing.

3. Experiments

(a) Apparatus

Stirred vessel: Figure 5.2 (Cylindrical flat bottom vessel, four baffles).
Impeller: FBDT impeller.

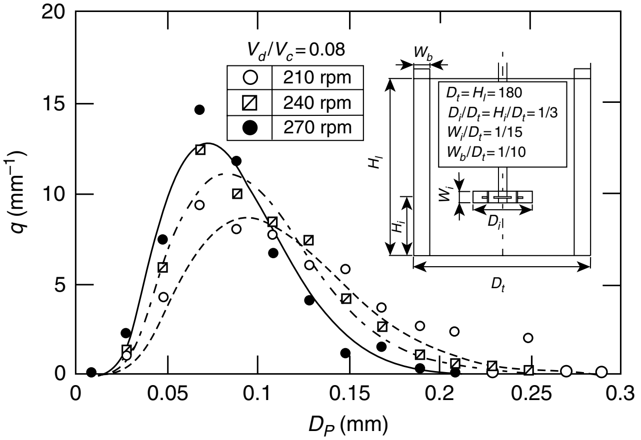


Figure 5.2 Data of droplet size probability density distribution in liquid–liquid mixing and fitted PSD curve based on new PSD function.

(b) Condition

System: water (continuous phase)–ethyl malonate (dispersion phase, 1 g/50 cc water (0.08 wt%), surface tension 11.2 mN/m, viscosity 2.0 mPa s, density 1.055 g/cm^3).

Impeller rotational speed (N): 210, 240, 270, 300 rpm ($Re = 1.26 \times 10^4$–1.8×10^4).

(c) Droplet measurement method

Camera (shutter speed 1/500 s).

(d) Procedure

The stirred vessel is placed in a square vessel filled with water in order to take pictures of the droplets using a camera. After confirming that the flow in the vessel attains a steady state under a fixed impeller rotational speed, images of the droplets are taken by a camera. By using these pictures, the droplet size distribution is estimated by fitting new PSD defined by Eq. (5.18).

4. Experimental result

Figure 5.2 (New PSD versus droplet size PSD).

5. Noteworthy point

(a) The droplet size distribution can be fitted by the new PSD defined by Eq. (5.18) irrespective of N.

(b) The PSD of the droplets exhibits a narrow distribution as N increases.

Challenge 5.3. Bubble size distribution in aerated stirred vessel

1. Scope

The aerated stirred vessel has been widely operated in chemical industries. The aim of gas–liquid mixing is to make the bubbles small in order to accelerate the mass transfer between the gas and the liquid. In the case of gas–liquid mixing, the gas that flows out from a sparger is trapped once by the impeller and then discharged as bubbles. The mixing state of the gas–liquid mixing is classified into three states:

(1) state controlled by mixing (refer to region with agitation effect in Challenge 2.4),
(2) state controlled by aeration (refer to region with agitation effect in Challenge 2.4, and analogical state of the bubble column),
(3) intermediate state of the two critical states (Figure 5.3).

The effect of the operation conditions on PSD in each state is different. Although the bubble size distribution in the operation has been expressed by various PSD functions, the PSD function that is most utilized is the normal PSD function. However, there is no physical background for applying the normal PSD function to the bubble size distribution. Additionally, when the bubble distribution is expressed by various PSD functions, it becomes difficult to discuss the relationship between the parameters in PSD and operation condition. This is one of the obstacles in the development of particle technology.

2. Aim

To examine whether the new PSD defined by Eq. (5.18) can sufficiently express bubble size distribution in an aerated stirred vessel.

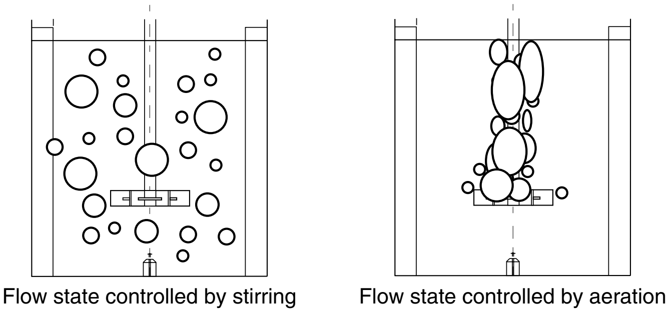


Figure 5.3 Flow states controlled by stirring and aeration.

3. Experiments

(a) Apparatus

Stirred vessel: Figure 5.4 (Cylindrical flat bottom vessel, four baffles).
Impeller: FBDT.
Nozzle: 3 mm ϕ (center of bottom).

(b) Condition

System: ion exchange water (liquid phase)–nitrogen gas (dispersion phase).
Impeller rotational speed (N): 150, 200, 250, 300, and 350 rpm ($Re = 0.90 \times 10^4$–2.1×10^4).
Gas flow rate: $1.67 \times 10^{-5}\ \mathrm{m^3/s}$.

(c) Measuring method of bubbles

Video camera.

(d) Procedure

The aerated stirred vessel is placed in a square vessel that is filled with water to obtain the picture of the bubbles. The impeller is rotated at a fixed rotational speed and the nitrogen gas is fed at a fixed flow rate. After confirming that the flow in the vessel has attained steady state, the images of bubbles are obtained by a video camera. By using these images, the bubble size distribution is estimated by fitting new PSD defined by Eq. (5.18).

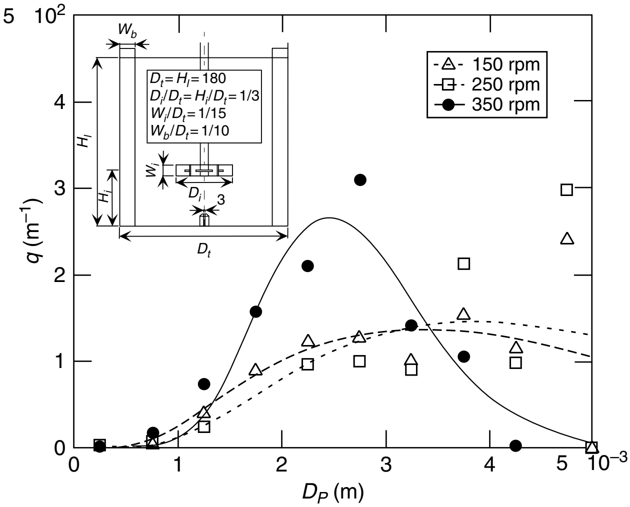


Figure 5.4 Data of bubble size probability density distribution in an aerated stirred vessel and fitted PSD curve based on new PSD function.

4. Experimental result

Figure 5.4 (A part of the result of the new PSD versus bubble size PSD).

5. Noteworthy point

(a) The bubble size distribution can be fitted by the new PSD defined by Eq. (5.18) irrespective of the values of N.

(b) If the impeller rotation speed is 150–250 rpm, the distribution has no peak and when N is 150–250 rpm, the distribution shows one peak. This can be understood as follows. With a large N, the large bubble stays for a long time in the vessel and is broken up by the impeller and shear stress. However, in the case of a low N, the gas injected flows out of the vessel without break up by the impeller and shear stress.

Challenge 5.4. Crystal size distribution

1. Scope

Recently, crystallization has been widely operated in chemical industries. Although the crystal size distribution in the operation has been expressed by using various PSD functions, the most utilized PSD function is the log-normal PSD function. However, there is no physical background for applying the log-normal PSD function to the crystal size distribution. Additionally, when the crystal size distribution is expressed by various PSD functions, it becomes difficult to discuss the relationship between the parameters in the PSD and the operation condition. This is one of the obstacles in the development of particle technology.

2. Aim

To examine whether the new PSD defined by Eq. (5.18) can sufficiently express the crystal size distribution in the crystallizer.

3. Experiments

(a) Apparatus

Crystallizer: Figure 5.5 (Cylindrical flat bottom vessel, four baffles).
Impeller: FBDT impeller.

(b) Condition

System: ion exchange water–potassium sulfate.
Impeller rotational speed (N): 500, 600, 700 rpm ($Re = 3.00 \times 10^4$–4.2×10^4).

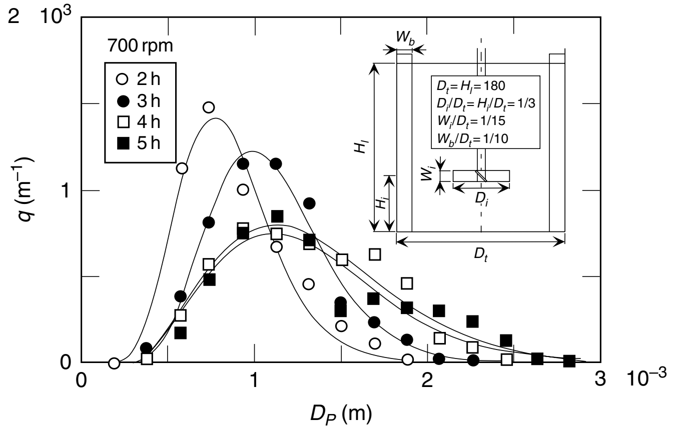


Figure 5.5 Data of crystal size probability density distribution and fitted PSD curve based on new PSD function.

(c) Measuring method of bubbles

Video camera.

(d) Procedure

First, both the solution vessel and the crystallizer are placed in a square vessel. Second, if the valve that connects both vessels is shut, each vessel is fixed at different temperatures (crystallizer: 20 °C, solution vessel: 20 °C) and concentrations (crystallizer: 105.9 kg/m^3, solution vessel: 136.5 kg/m^3).

Third, the valve that connects both vessels is open and about 1 mg of the mother crystal is fed after confirming that the temperature in both vessels has become constant (after about 1 h). After 2 h, the images of crystals are taken by a video camera under irradiation through a 4 mm slit every 1 h after adding the mother crystal. Based on these pictures, the crystal size distribution is estimated by fitting new PSD defined by Eq. (5.18).

4. Experimental result

Figure 5.5 (A part of the result of new PSD versus crystal size PSD).

5. Noteworthy point

(a) The crystal size distribution can be fitted by the new PSD defined by Eq. (5.18) irrespective of the lapse of time.

(b) Four hours after adding the mother crystal, the crystal size distribution becomes a constant distribution.

Challenge 5.5. Crushed product size distribution

1. Scope

Recently, crushing has been widely used in chemical industries. Although the crushed product size distribution in the operation has been expressed by various PSD functions, the most utilized PSD function is the log-normal PSD function. However, there is no physical background for applying the log-normal PSD function to the crushed product size distribution. Additionally, when the crushed product size distribution is expressed by various PSD functions, it becomes difficult to discuss the relationship between the parameters in the PSD and the operation condition. This fact is one of the obstacles in the development of particle technology.

2. Aim

To examine whether the new PSD defined by Eq. (5.18) can sufficiently express the crushed product size distribution in the grinder.

3. Experiments

(a) Apparatus

Crusher: ball mill ($156\,\text{mm}\phi \times 156\,\text{mm}$), ceramic ball ($25\,\text{mm}\phi \times 30$).

(b) Condition

Raw feed material: constant diameter limestone (density $2.53 \times 10^3\,\text{kg/m}^3$ and $2.65 \times 10^3\,\text{kg/m}^3$, classified by Tyler standard sieve of $200\,\text{mm}\phi$).
Ball mill revolution speed: 50, 100, and 150 rpm
Feed material: $D_o = 1.086,\ 1.524,\ 1.816\,\text{mm}$ for a density of $2.53 \times 10^3\,\text{kg/m}^3$, $D_o = 0.456,\ 0.645,\ 1.283,\ 1.524\,\text{mm}$ for $2.65 \times 10^3\,\text{kg/m}^3$
Weight of feed material: 600 g

(c) Procedure

A fixed weight feed is fed in the ball mill and the ball mill is rotated at a fixed rotational speed. After the rotation has begun, the crushed product is removed at constant time interval, and the crushed product size distribution is measured by using a Tyler standard sieve.

4. Experimental result

Figure 5.6 (A part of the result of new PSD versus crushed product size PSD).

5. Noteworthy point

(a) Even if limestone with a constant diameter is crushed, the crushed product size distribution can be sufficiently fitted by the new PSD expressed by Eq. (5.18) regardless of the time.

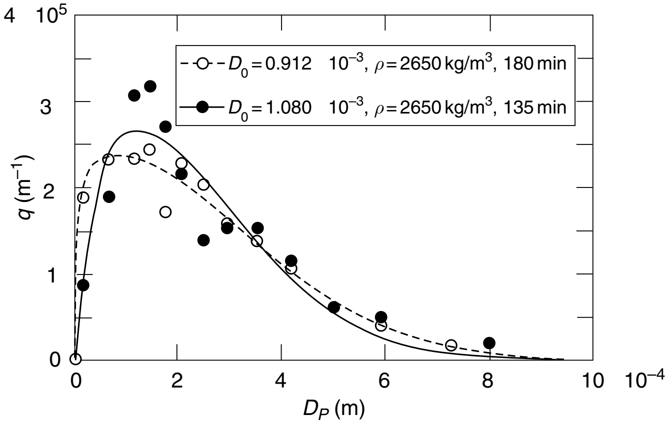


Figure 5.6 Data of crushed product size probability density distribution and fitted PSD curve based on new PSD function.

5.4 Summary

In this chapter, the following points have been clarified:

(1) The general expression of the PSD is presented by making use of information entropy. It is confirmed that this newly defined expression can effectively represent both narrow and wide distributions that are described by three traditional common PSDs.

(2) The usefulness of the new PSD is clarified by applying it to the following examples:

   (a) the superiority of the new PSD in comparison to traditional PSD with regard to usefulness,
   (b) droplet size distribution in liquid–liquid mixing in a stirred vessel,
   (c) bubble size distribution in an aerated stirred vessel,
   (d) crystal size distribution in a crystallizer,
   (e) crushed product size distribution in a ball mill.

(3) The eddy size probability density distribution can be expressed by the newly defined PSD.

CHAPTER 6

Anxiety/Expectation

6.1 Introduction

In the previous chapters, the apparently appropriate field of chemical engineering was examined from the viewpoint of information entropy; further, newly obtained knowledge was described. Our surroundings are rapidly changing with the developments in science and technology. Already, safety and a sense of security cannot be obtained without consciousness, that is, consciousness and investment have become indispensable. The expectations from science and technology for constructing a peaceful society are very huge. In order to solve the problem related to the peace of mind at a higher level, it is necessary to consider the safety provided by science and technology and the need for a sense of security together. Therefore, many sciences must work together. The inclusion of psychology, which regards the sense of security and the probability terms, could be one such joint initiative.

In this section, we change our viewpoint slightly from the pure chemical engineering field; we now focus on the anxiety/expectation that deeply affects human decision-making. The reader might believe that safety is excusable but the anxiety/expectation is far from reality in chemical engineering. However, a chemical engineer studies almost all the problems surrounding human welfare and experiences anxiety/expectation daily. Additionally, there is a possibility that he/she makes a decision with regard to matters in the field of chemical engineering. The decision-making, as described above, can be considered to be affected by anxiety/expectation with regard to the phenomena that might be a result from the decision. Therefore, the decision-making is equalized to discuss about the anxiety/expectation. Taking into consideration the discussion above, it will not be a misdirected policy to give anxiety/expectation greater importance and discuss decision-making. By making anxiety the target of study, chemical engineering will have a new approach—as a study of unification. By considering our surroundings as described above, let us now shift our focus from pure chemical engineering to anxiety/expectation.

However, there is no established quantitative expression for the degree of anxiety/expectation. Further, it is indispensable to discuss a method of defining the degree of anxiety/expectation. Anxiety/expectation consists of the value of the object of the phenomenon and the probability of occurrence of the phenomenon. In the traditional simple way of considering decision-making, the product of the value and probability of the occurrence of the result are important parts of decision-making. However, it is clarified that the traditional way of thinking about the decision-making is not perfect and has many limitations. Recently, many modern methods of considering the decision-making have been suggested and analytic hierarchy process (AHP) is considered to be an ultramodern method of decision-making. However, there is an unreasonable point about this method as well. This method includes a step to determine the strengths or the priorities of the phenomena/elements based on pairwise comparison. The pairwise comparison involves the comparison of a phenomenon/element of one level of a hierarchy with that of the next level. In more concrete terms, for example, for given phenomena/elements A and B, the series of pairwise comparison is determined in the case of five judgments in the following manner:

(1) If A and B are equally important, insert 1.
(2) If A is weakly more important than B, insert 3.
(3) If A is strongly more important than B, insert 5.
(4) If A is demonstrably or very strongly more important than B, insert 7.
(5) If A is absolutely more important than B, insert 9.

Here, the numbers that are to be inserted are referred to as the "agreed-upon" number, and there is no physical background to determine their value. The decision-making based on the "agreed-upon" numbers is very qualitative, because the five judgments and the agreed-upon numbers in the above example are determined by human experience. The agreed-upon number will indicate the meaning of the value of the phenomenon caused by the decision. Of course, there is a possibility that the number includes the probability of occurrence of the object of the phenomenon. However, the pairwise comparison—that is, the "agreed-upon" number—establishes the priorities of the elements of one level of a hierarchy with respect to one element of the next level. If there is psychological motivation or any such thing, the method will be clearer but currently, this is an impossible demand. Nevertheless, a human being makes a decision quantitatively and unintentionally. When a human makes a decision, a comparison is made between the phenomena/elements based on various evaluation standards, and the probability of occurrence and value of the phenomena/elements as a result of the decision are not fully considered. However, as observed in the case of AHP, the probability of occurrence of the phenomenon/element as a result of the decision is not sufficiently considered when decision-making is discussed. Therefore, it is very important to build a new line of thought that has significant meaning in the decision-making based on the probability of occurrence of the phenomena/elements due to the decision. Henceforth, information entropy is considered to have a possibility to present a new way of thinking. This is because human experience for quantity has a close relation with the information entropy, and there is a possibility that a change in human experience can be expressed by the information entropy, as shown in Chapter 1. In this chapter, the usefulness of information entropy over the sensation of anxiety/expectation is investigated. It is expected that in the future, this field will be certainly included in the field of chemical engineering.

Let us deal with AHP. This method was developed by Saaty in 1971 for the decision-making method of problems under uncertain conditions and various evaluation standards. The most important point of this method is that the structure of the decision-making is considered to be hierarchical. The basis of the calculation procedure of decision-making is as follows:

(1) According to the analysis of the problem, the format for hierarchies and decomposition is set.
(2) By making use of pairwise comparison between an element in one level and another in the next level, a matrix of the evaluation terms is created.
(3) If the ratio of C.I. (consistency index) to R.I. (random index)—C.R. (= C.I./R.I.)—has an unreasonable value, the above step is repeated.

Based on the result of the pairwise comparison, the synthesized weight of the elements is calculated and the total evaluation value is determined.

6.2 Safety and anxiety

"Safety" has always been discussed in various fields. The opposite of safety is danger. In general, the security for an industrial equipment/process/plant, even in the case of those rated as perfectly safe, is rarely felt to be perfect by the common man. A good example of such a situation is nuclear power stations. Although nuclear power stations are considered to be technologically perfect in almost all aspects, the installation of a new plant is not easily accepted by the residents in and around the site. This is because the residents do not feel perfectly secure; instead, they are anxious about the occurrence of a serious accident, even though the chances of such an occurrence are rather small.

With regard to the anxiety for physical damages, "safety" is equivalent to "security" and "danger" to "anxiety." However, in the other cases, "safety" does not always mean "security." Safety equipment, safety process, and safety plant imply harmless equipment, harmless process, and harmless plant, respectively. On the other hand, secure equipment, secure process, and secure plant should not only be harmless but should also guarantee peace of mind.

Although the degree of danger generally increases in proportion to the probability of an undesirable event, the degree of anxiety will not be proportional to the probability. The degree of anxiety changes with the probability, giving an S-shaped curve that is convex and concave for small and large values of probability, respectively. In other words, when an undesirable instance that has previously been believed to be impossible becomes possible, people suddenly feel anxious. The degree of anxiety increases rapidly depending on the increase in the probability. However, in the case of moderate probabilities, the degree of anxiety increases gradually. From a certain value of the probability, the anxiety curve again shows a rapid increase because possibility is close to changing into certainty. In other words, there is a region of moderate increase in the change in the degree of anxiety with probability. It is true that security is required in decision-making, for example, whether to adopt the means to improve or whether to install new equipment/process/equipment. However, there is no suitable index to quantitatively measure the degree of anxiety. A new index to quantitatively express the degree of anxiety/expectation, under the condition that the probability of occurrence of an undesirable/desirable event is known, is discussed.

6.3 Evaluation index of anxiety/expectation[2]

(1) Discussed on anxiety

People experience anxiety/expectation when there is a possibility that an undesirable/desirable event will occur. Anxiety/expectation can be discussed by using the following five items:

1. Type of situation: (a) hard type and (b) soft type
2. Authorization background: (a) scientifically guaranteed level, (b) approved level by an individual, (c) approved level by a certain group, (d) realizable level, (e) realized level, and (f) ideal level
3. Scale of impact: (a) individual, (b) group, (c) society, and (d) environment
4. Object of damage: (a) material, (b) physical, (c) social, and (d) mental
5. Other information: (a) condition, (b) probability, (c) scale, and (d) strength of the damage

As described above, the core of anxiety/expectation is very complex, and it has been very difficult to quantitatively measure the degree of anxiety/expectation. In the following section, a method of expressing the degree of anxiety/expectation is discussed regardless of the items described above (type of situation, authorization background, scale of impact, and object of damage) under the condition where the probability of the occurrence/disappearance and weight of the value of an undesirable/desirable situation, that is, P and V in Eq. (6.7) below, are known.

(2) Definition of the degree of anxiety/expectation

It is natural to consider that the degree of anxiety/expectation is determined by the probability of occurrence and the weight of the value of the undesirable/desirable event. The degree of anxiety/expectation strongly depends on the uncertainty regarding the occurrence/disappearance of the undesirable/desirable event. Security, which is the opposite of anxiety, is installed in locations where there is no anxiety. On the other hand, the meaning of expectation becomes identical to that of anxiety when the way it is considered is inverted. In other words, the relation between the degree of anxiety and that of expectation is contrary, and the degree of expectation can be defined similar to the degree of anxiety. In the following discussion, only the degree of anxiety with respect to the occurrence of the undesirable event is considered, for simplicity.

The anxiety is due to the existence of uncertainty regarding "whether such an undesirable event occurs." The uncertainty is characterized by the probability of occurrence, P, and the probability of disappearance, $(1-P)$. The amount of uncertainty regarding whether the undesirable event occurs is expressed by the information entropy as

$$H = -P\ln P - (1-P)\ln(1-P) \tag{6.1}$$

The relationship between H and P is shown in Figure 6.1.

H has the maximum value of ln 2 at $P = 1/2$, and the distribution of H becomes symmetric with respect to $P = 1/2$. By considering that $(1-P)$ implies the probability of disappearance of undesirable event, the amount of uncertainty becomes maximum when the probability of occurrence and probability of disappearance are equal. For the probabilities, there are two natural reference points—

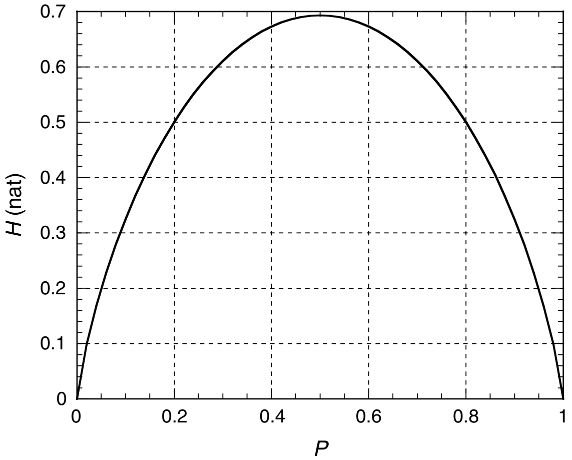


Figure 6.1 Information entropy distribution.

impossibility and certainty. The two points correspond to the end points of the scale when the degree of anxiety is discussed, because it is natural to consider that the anxiety regarding the occurrence of the undesirable event is deeply related to the amount of uncertainty. In particular, it can be considered that the change in the degree of anxiety has a significant connection with the change in the amount of information entropy. The difference between the maximum value of the amount of uncertainty $H_{\max}$ and H has a very important role in this discussion:

$$H_{\max} - H = \ln 2 - \{-P \ln P - (1 - P)\ln(1 - P)\} \tag{6.2}$$

The relationship between $(H_{\max} - H)$ and P is shown in Figure 6.2.

$(H_{\max} - H)$, which is the decrease in the amount of uncertainty regarding the point whether the undesirable event occurs, can be understood as follows according to the range in the value of the probability of occurrence. Besides, in the following discussion, let CO be the amount of certainty regarding the occurrence of the undesirable event and let CD be the amount of certainty regarding the disappearance of the undesirable event.

$$H_{\max} - H = 0 : \mathrm{CO} = \mathrm{CD} \tag{6.3a}$$

$$H_{\max} - H \neq 0 : \mathrm{CO} \neq \mathrm{CD}$$

$$P < 1/2 : \mathrm{CO} < \mathrm{CD} \text{ and } H_{\max} - H = \mathrm{CD} - \mathrm{CO} = -(\mathrm{CO} - \mathrm{CD}) \tag{6.3b}$$

$$P > 1/2 : \mathrm{CO} > \mathrm{CD} \text{ and } H_{\max} - H = \mathrm{CO} - \mathrm{CD} \tag{6.3c}$$

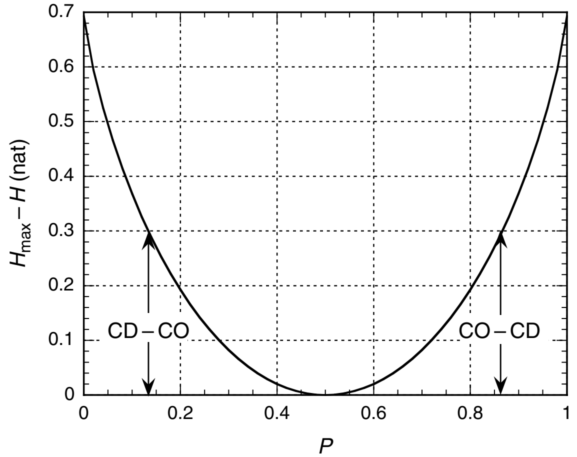


Figure 6.2 Difference between maximum amount of information entropy and amount of information entropy at arbitrary probability value.

When the anxiety regarding the occurrence of an undesirable event is discussed, the reference point that is to be considered should be consistent. By considering that the amount of anxiety regarding the occurrence of the undesirable event increases with an increase in the certainty regarding the occurrence and decrease in the certainty regarding the disappearance of the undesirable event, it is assumed that the degree of anxiety regarding the occurrence of the undesirable event is proportional to the specific value that is based on the value of (CO–CD) as

$$P < 1/2: \text{ the degree of anxiety } \mathrm{CO} - \mathrm{CD} = -(H_{\max} - H) \tag{6.4a}$$

$$P > 1/2: \text{ the degree of anxiety } \mathrm{CO} - \mathrm{CD} = (H_{\max} - H) \tag{6.4b}$$

From the above consideration, the degree of anxiety should be discussed by using negative values when $P < 1/2$; further, it has a minimum value that corresponds to $-H_{\max}$. Let us consider that the degree of anxiety regarding the occurrence of an undesirable event should be positive and $P = 0$ should be considered as the origin for the anxiety regarding the occurrence of an undesirable event. In such a case, the above difference in information entropy should be treated by shifting the origin on the positive side by $H_{\max}$ as

$$\Delta I_{P<1/2} = (H_{\max}) - (H_{\max} - H) = H \tag{6.5a}$$

$$\Delta I_{P\geq 1/2} = (H_{\max}) + (H_{\max} - H) = 2H_{\max} - H \tag{6.5b}$$

According to the assumption that the degree of anxiety regarding the occurrence of the undesirable event, AE, is proportional to the change in the amount of information entropy described above, the following relationships can be obtained:

$$\mathrm{AE}_{P<1/2} \propto \Delta I_{P<1/2} = (H_{\max}) - (H_{\max} - H) = H \tag{6.6a}$$

$$\mathrm{AE}_{P\geq 1/2} \propto \Delta I_{P\geq 1/2} = (H_{\max}) + (H_{\max} - H) = 2H_{\max} - H \tag{6.6b}$$

Sufficient attention must be focused on the fact that there is no factor of human experience in the development; it then becomes possible to define the degree of anxiety based only on the change in the amount of information entropy regarding the appearance and disappearance of the undesirable event. In other words, Eq. (6.5) should not be recognized only as the amount of uncertainty regarding the appearance and disappearance of the undesirable event.

The weight of the value of the undesirable event is not constant and is due to each undesirable event. When the weight of the value of the undesirable event is V (an arbitrary unit such as \$), the degree of anxiety regarding the occurrence of the undesirable event can be defined by the following equations:

$$\mathrm{AE}_{P<1/2} = V\{-P\ln P - (1-P)\ln(1-P)\} \tag{6.7a}$$

$$\mathrm{AE}_{P\geq 1/2} = V[2\ln 2 - \{-P\ln P - (1-P)\ln(1-P)\}] \tag{6.7b}$$

In the following discussion, V is termed the value-factor. The relationship between AE and P when $V = 1$ is shown in Figure 6.3.

From this figure, the relationship gives an S-shaped curve and it becomes clear that the degree of anxiety overweighs low probabilities and underweighs high probabilities. The S-shaped curve is similar to the relationship between the decision weight and stated probability by Tversky and Fox[1]. Therefore, the S-shaped curve implies that increasing the probability of occurrence by 0.1 has a greater impact when it changes the probability of occurrence from 0.9 to 1.0 or from 0 to 0.1 than when it changes the probability of occurrence from, say, 0.3 to 0.4 or from 0.6 to 0.7. This degree of anxiety has a minimum value of zero at $P = 0$ and a maximum value of $2V$ ln 2 at $P = 1$. The maximum value depends on the value of V, that is, the maximum value is decided by the value-factor of the undesirable event. However, the standardized distribution of the degree of anxiety by using its maximum value, $\mathrm{AE}_{P=1}$, becomes similar regardless of the value-factor of the undesirable event.

As mentioned before, the expression of the definition of the degree of expectation becomes identical to that of anxiety, Eq. (6.7), because only the viewpoint regarding the probability becomes opposite to that of the definition of the degree of anxiety.

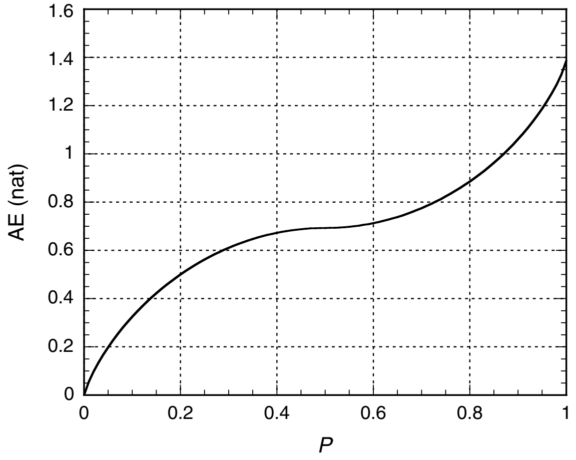


Figure 6.3 Anxiety/expectation–probability curve.

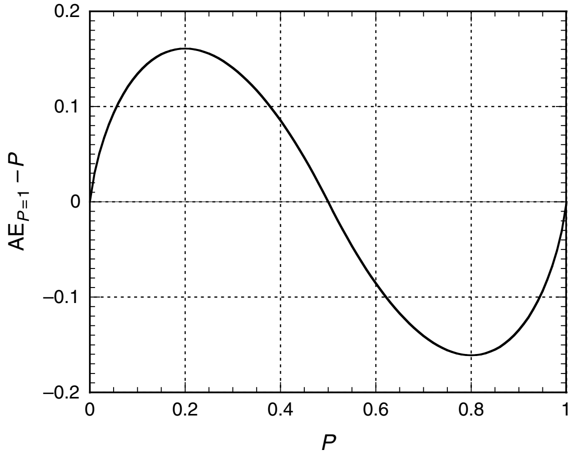


Figure 6.4 Difference between objective probability and subjective probability.

When the probability of occurrence, P, described above appears to be the "objective probability (stated probability)," the degree of anxiety that is standardized by using its maximum value, $AE_{P=1}$, can be regarded as the "subjective probability." From this consideration, it is possible to explain the above new degree of anxiety as follows. There is a two-stage process in which people first assess the objective probability of an uncertain event and then transform this value by a subjective probability that is in proportion to the degree of anxiety. The relationship between the difference of the two probabilities and objective probability is shown in Figure 6.4.

6.4 Utilization method and usefulness of newly defined degree of anxiety

(1) Anxiety regarding occurrence of accident

There can be many major accidents around us—nuclear power station accidents, general plant accidents, traffic accidents, and so on. Anxiety associated with these accidents is considered here. If the object of anxiety is physical, the accident situations can be classified into the following three stages:

(1) Whether an injury has been suffered?
(2) Whether the suffered injury is serious?
(3) Whether the serious injury results in death?

It is possible to express the degree of anxiety regarding each stage by using Eq. (6.7). If there is no information regarding whether the accident reaches the final stage, it is natural to discuss the anxiety in the stages described above, according to the nature of the accident. A new installation of a nuclear power station is not accepted easily at the site because the value-factor of the power plant accident is fairly large and the degree of anxiety regarding the accident increases more than that of another type of accident, even if the probability of occurrence of a nuclear power station accident is rather small.

Challenge 6.1. Anxiety regarding involvement in an accident in outdoors

1. Scope

When we go outdoors, we are worried about our safety. Although the anxiety regarding the occurrence of an accident outdoors is estimated by using statistical data and so on, there are few reports that discuss the degree of anxiety regarding involvement in such accidents when someone is about to leave the place. Therefore, an examination of whether the new degree of anxiety can express the appropriate degree of anxiety is indispensable.

2. Aim

To examine whether the new degree of anxiety defined by Eq. (6.7) can express the appropriate degree of anxiety.

3. Calculation

(a) Condition

Steps for an accident: three steps (“Whether an injury has been suffered?,” “Whether the suffered injury is serious?,” and “Whether the serious injury results in death?”).

Value and estimated probability of occurrence: Stage 1: $V = 1$ and $P = 0.8$, Stage 2: $V = 4$ and $P = 0.5$, Stage 3: $V = 10$.

(b) Method

The distribution of anxiety at the second stage begins from the amount of the degree of anxiety at the previous stage. Therefore, the degree of anxiety when $P = 0$ of stage 2 is the same as at $P = 0.8$ of stage 1, and the degree of anxiety at $P = 0$ of stage 3 is the same value at $P = 0.5$ of stage 2. The degree of anxiety at each stage can be calculated by using Eq. (6.7).

4. Calculated result

Figure 6.5 (Anxiety–probability curve).

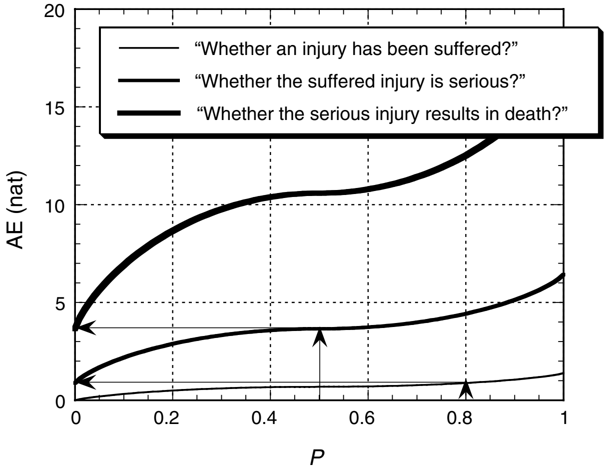


Figure 6.5 Anxiety–probability curve in the case of accident in outdoors.

5. Noteworthy point

(a) The new degree of anxiety can sufficiently express the appropriate degree of anxiety of being involved in the accident at the time of leaving.

6. Supplementary point

(a) The degree of anxiety at each stage changes with change in the value-factor at each stage.

(2) Decision-making of significant issues

There are various kinds of decision-making processes in the chemical industry field. An equipment/process/plant is improved or a new process/plant is built without definite knowledge of their consequences. When the definition of the degree of anxiety expressed in Eq. (6.7) is applied to decision-making, it sometimes becomes necessary to change the considerations regarding object and probability from anxiety to expectation and from the probability of occurrence of an undesirable event to the probability of success of desirable event.

(a) *How to prioritize the units for improvement*

Often, the units to be improved must be prioritized. For this purpose, the following case is considered. There are multiple units that require improvement. The value-factor and probability of occurrence of the accident/trouble of each unit are known. In this case, it is possible to prioritize improvements among units based on the order of the amount of degree of anxiety regarding the occurrence of the accident/trouble. The unit that has the maximum degree of anxiety has the highest priority for improvement. It is necessary to decide the value-factor for each unit by considering all the possible properties that can be caught by the decision-maker. Of course, in this case, it does not matter much whether the value-factor is expressed as the relative value based on the value of the standard state or not.

Challenge 6.2. Priority between two units for improvement

1. Scope

In many cases, multiple units must be improved. In such cases, it is necessary to decide the priority among them for improvement. It can be stated that decision-making has been done either from intuition or a very simple comparison among the units. This is one of the methods of deciding the priority for improvement among them in the order of the amount of degree of anxiety. Therefore, examining whether the new degree of anxiety can decide the priority among the units to be improved is indispensable.

2. Aim

To examine whether the new degree of anxiety defined by Eq. (6.7) can decide the priority among the units for improvement.

3. Calculation

(a) Condition

Number of units: two units (Unit 1 and Unit 2).
Value of each unit: $V1 = 2$ and $V2 = 1$.
Probability of occurrence of the trouble: $P1 = 0.2$ and $P2 = 0.8$.

(b) Procedure

The degree of anxiety regarding each unit is calculated based on Eq. (6.7).

4. Calculated result

Figure 6.6 (Anxiety–probability curve).

5. Noteworthy point

(a) When the degree of anxiety is assumed to be simply proportional to the probability of occurrence of the trouble and the value-factor is the same as the earlier one, the degree of anxiety of Unit 1 and Unit 2 is expressed as 0.4 and 0.8, respectively. According to this earlier consideration, Unit 2 has a higher priority for improvement than Unit 1. However, when the newly defined degree of anxiety that is expressed in Eq. (6.7) is applied,

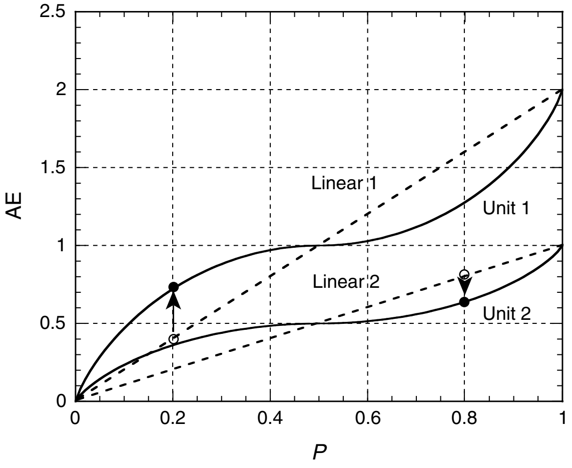


Figure 6.6 Priority between two units to improve.

the relationship between the degree of anxiety and the probability (anxiety–probability curve) shows a contrary result. In other words, Unit 1 has a higher priority for improvement than Unit 2 because the degrees of anxiety are 0.72 for Unit 1 and 0.64 for Unit 2.

(b) *Whether the means of improvement must be adopted*

Often, a means of improvement must be adopted. The following cases are considered.

(i) Improvement for increasing profit

An equipment/process makes a profit A. When the means for improving the equipment/process is adopted, the accomplishment makes an amount of profit of B $(B > A)$. The probability of success of the improvement is known as P. Under these conditions, the decision-making with regard to whether the means should be adopted is discussed. In order to make a decision, the relationship between the degree of expectation and probability of success is plotted under the condition that the value-factor takes the value of B that is achieved when the means succeeds in improving. When the degree of expectation is larger than the profit at present A, the means should be adopted. On the other hand, when the degree of expectation is equal to or smaller than A, the means should not be adopted. It is necessary to decide the profit by considering all the properties that can be decided by a decision-maker. It does not matter much to express the profit as a relative value based on the value of the standard state. This line of thinking can be used to choose the most secure means to improve process/equipment in multiple means.

Challenge 6.3. Whether the proposal to improve the unit must be adopted

1. Scope

There are several instances when an administrator must decide whether the proposal of improvement regarding the unit or plant and other aspects should be adopted. There is no clear standard for taking decisions on this. It can be stated with some fairness that the decision-making has been a result of intuition or a very simple comparison among units and plants. This is one of the ways of deciding whether the proposal of improvement should be adopted in the order of the amount of degree of expectation. Therefore, it is indispensable to examine whether the proposal of improvement should be adopted.

2. Aim

To examine whether the proposal for improvement should be adopted based on the new degree of expectation defined by Eq. (6.7).

3. Calculation

(a) Condition

Gain at present (A): 400 million dollars.
Gain when improvement succeeded (B): 1000 million dollars.

(b) Method

The degree of expectation of the success of improvement is calculated based on Eq. (6.7) under the condition that the value-factor takes the value of \$1000 million.

4. Calculated result

Figure 6.7 (Expectation–probability curve).

5. Noteworthy point

(a) The proposal should be adopted when $P > 0.24$, and the means should not be adopted when $P \leq 0.24$.

(ii) Improvement for decreasing loss

An equipment/process makes a loss A. When the means is adopted to improve the equipment/process, the accomplishment does not make a loss. On the other hand, when the means fails, the failure makes a loss B $(B > A)$. The probability of failure of the improvement is known as P. Under these conditions, the decision-making with regard to whether the means should be adopted is discussed. In order to make the decision, the relationship between the degree of anxiety and the probability of failure is plotted under the condition that the value-factor has a value of B that is attained when the means fails to improve. When the degree

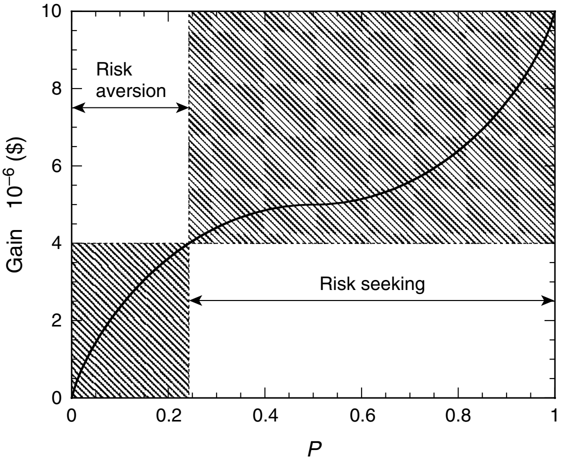


Figure 6.7 Expectation–probability curve for decision-making to adopt the means to improve.

of anxiety is equal to or larger than the present loss A, the means should not be adopted. On the other hand, when the degree of anxiety is less than A, the means should be adopted. It is necessary to decide the loss by considering all the possible properties that can be decided by the decision-maker. Of course, expressing the loss as the relative value based on the value of the standard state matters very little. This line of thinking can be used to choose the most secure means to improve the process/equipment in multiple ways.

Challenge 6.4. Whether the means for improvement should be adopted

1. Scope

There are several instances when an administrator must decide whether the proposal of improvement regarding units or plants and other aspects should be adopted. There is no clear standard to make a decision concerning this matter. It can be stated with some fairness that the decision-making has been done by human experience or by a very simple comparison among units and plants. This is one of the ways to decide whether the proposal for improvement should be adopted based on the amount of degree of anxiety. Therefore, examining whether the proposal for improvement should be adopted is indispensable.

2. Aim

To examine whether the proposal for improvement should be adopted on the basis of the new degree of anxiety defined by Eq. (6.7).

3. Calculation

(a) Condition

Loss at present (A): \$600 million.
Loss when improvement failed (B): \$1000 million.

(b) Method

The degree of anxiety of the success of the improvement is calculated based on Eq. (6.7) under the condition that the value-factor takes the value of \$1000 million.

4. Calculated result

Figure 6.8 (Anxiety–probability curve).

5. Noteworthy point

(a) The proposal should not be adopted when $P \geq 0.76$, and the means should be adopted when $P < 0.76$.

(c) *Whether the new equipment/process must be installed*

It is often necessary to install a new equipment/process. The following case is considered. The installation of a new equipment/process by a certain method incurs a cost A. When the installation is successful, it makes a profit B ($B > A$). The probability of the installation being successful is P. Under these conditions, the decision-making regarding whether the new equipment/process should be installed is discussed. In order to take the decision, the relationship between the

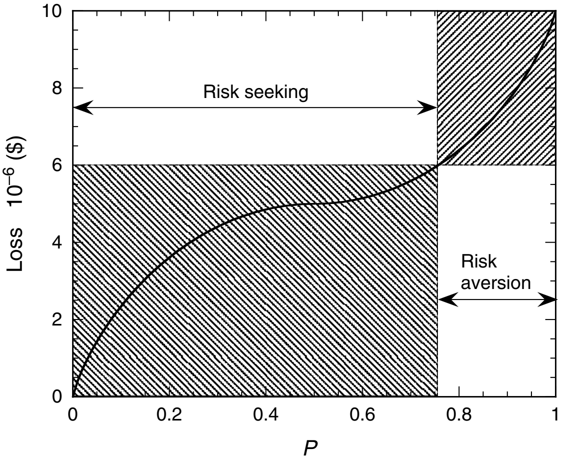


Figure 6.8 Anxiety–probability curve for decision-making to adopt the means to improve.

degree of expectation and probability of success is plotted under the condition that the value-factor takes the value B that is attained when the installation is successful. If the degree of expectation is larger than the construction cost A, the new process/equipment should be installed. On the other hand, if the degree of expectation is equal to or smaller than A, the new process/equipment should not be installed. The relation described above can be explained by using Figure 6.3. It is necessary to decide the installation cost and profit by considering all of the possible properties that can be decided by the decision-maker. Of course, expressing the installation cost and profit as relative values based on the value of the standard state is not very important. This line of thinking can be used to choose the most secure process/equipment in multiple ways in new projects/plants.

It is clarified from Figure 6.3 that when the installation cost is half the profit after a successful installation ($A = B/2$), decision-making in the region of the probability of success of $0.4 < P < 0.6$ is not easy; this is because the slope of the expectation–probability curve is quite small in this region. On the other hand, it is easy to make a decision in the region of the probability of success of $P < 0.4$ or $P > 0.6$ because the slope of the expectation–probability curve is clearly large in this region.

Challenge 6.5. Collation with old discussions

1. Scope

Many experimental data concerning decision-making exist. Therefore, it is indispensable to collate such data, and the result that is obtained by using a new degree of anxiety/expectation.

2. Aim

To collate the result that is obtained by using a new degree of anxiety/expectation defined by Eq. (6.7) with data reported by Tversky and Fox[1].

3. Calculation

(a) Condition

Fourfold pattern: Tversky and Fox[1] showed a common pattern of risk seeking and risk aversion observed in choices between simple prospects as follows. "$C(x, P)$ is the median certainty equivalent of the prospect (x, P). The Fourfold (a) in Table 6.1 shows that the median participant is indifferent between receiving a certain \$14 and a 5% chance of receiving \$100. Because the expected value of this prospect is only \$5, this observation reflects risk seeking. The Fourfold (a)–(e) in Table 6.1 illustrates a fourfold pattern of risk attitude: risk seeking for gains and risk aversion

Table 6.1 Fourfold pattern and winning results.

	Tversky and Fox $C(x, P)$: Median Certainty Equivalent of Prospect (x, P)	Author (Based on New Equation)
Fourfold		
(a)	C($100, 0.05) = $14	($100, 0.05) = $14
(b)	C($100, 0.95) = $78	($100, 0.95) = $84
(c)	C(−$100, 0.05) = −$8	(−$100, 0.05) = −$14
(d)	C(−$100, 0.95) = −$84	(−$100, 0.95) = −$84
Winning		
(a)	($30, 1.0) > ($45, 0.80)	($30, 1.0) = $30 > ($45, 0.80) = 29
(b)	($45, 0.20) > ($30, 0.25)	($45, 0.20) = $16 > ($30, 0.25) = $12
(c)	($100, 1.0) > ($200, 0.50)	($100, 1.0) = $100 = ($200, 0.50) = $100

(x, P): probability P chance of receiving x.

for losses with low probability coupled with risk aversion for gains and risk seeking for losses with high probability."

Risk seeking is exhibited if a prospect is preferred to a sure outcome with an equal or greater expected value. On the other hand, risk aversion is defined as a preference for a sure outcome over a prospect with an equal or greater expected value.

Winning:

(1) Tversky and Fox[1] wrote that most people prefer a certain $30 to an 80% chance of winning $45; further, most people also prefer a 20% chance of winning $45 to 25% chance of winning $30 as shown by Winning (a) and (b) in Table 6.1.

(2) Tversky and Fox[1] also wrote that a definite amount of $100 is chosen over an even chance to win $200 or nothing as shown by Winning (c) in Table 6.1.

Betting: Tversky and Fox[1] said that there is a pattern that violates the expected utility theory as follows: "112 Stanford students were asked to choose between prospects defined by the outcome of an upcoming football game between Stanford and the University of California at Berkeley. Each participant was presented with three pairs of prospects displayed in Table 6.2. The percentage of respondents who chose each prospect appears on the right-hand side."

"Table 6.2 shows that, overall, f_1 was chosen over g_1, f_2 over g_2, and g_3 over f_3. Furthermore, the triple (f_1, f_2, g_3) was the single most common pattern, selected by 36% of the respondents. This pattern violates expected utility theory, which implies that a person who chooses f_1 over g_1 and f_2 over g_2 should also choose f_3 over g_3. However, 64% of the

Table 6.2 Percentage of respondents of betting.

Problem	Option	Events				Preference (%)
		A($)	B($)	C($)	D($)	
1	f_1	25	0	0	0	61
	g_1	0	0	10	10	39
2	f_2	0	0	0	25	66
	g_2	10	10	0	0	34
3	f_3	25	0	0	25	29
	g_3	10	10	10	10	71

A: Stanford wins by 7 or more points; B: Stanford wins by less than 7 points; C: Berkley ties or wins by less than 7 points; D: Berkley wins by 7 or more points; Preference: percentage of respondents that chose each option (Tversky and Fox).

55 participants who chose f_1 and f_2 in Problems 1 and 2 chose g_3 in Problem 3, contrary to the expected theory."

(b) Method

According to the degree of anxiety/expectation defined by Eq. (6.7) the degree of anxiety/expectation is calculated under the given conditions.

4. Experimental result

Fourfold pattern: According to the new degree of anxiety/expectation defined by Eq. (6.7), the border between risk seeking and risk aversion becomes the same regardless of the gain and loss as

$$C(\$100,\ 0.05) = \$14 \quad C(-\$100,\ 0.05) = -\$14$$
$$C(\$100,\ 0.95) = \$84 \quad C(-\$100,\ 0.95) = -\$84$$

However, although there is a small difference between the results given by the author and by Tversky and Fox[1], in the case of loss (risk aversion), it can be said that the difference between them is not significant.

Winning:

(1) Based on the newly defined index, the case of 80% chance of winning \$45 corresponds to winning \$29, and this value is lower than \$30. Additionally, based on the newly defined index, the case of 20% chance of winning \$45 corresponds to winning \$16, and the case of 25% chance of winning \$30 corresponds to winning \$12; this value is lower than \$16. These results are shown in Table 6.1. In other words, the preference by the people described above can be clearly explained by the newly defined index.

(2) According to the new degree of anxiety/expectation in Eq. (6.7), an even chance to win \$200 or nothing corresponds to winning \$100, and there is no change between them as shown in Table 6.1. It is considerable that this case is controlled by intuitive reasoning.

Betting: This pattern, however, is consistent with the present new degree of anxiety/expectation in Eq. (6.7). The relationship between the probability of the occurrence of an option and the expectation to outcome is shown in Figure 6.9.

Let us deal with expected utility theory.

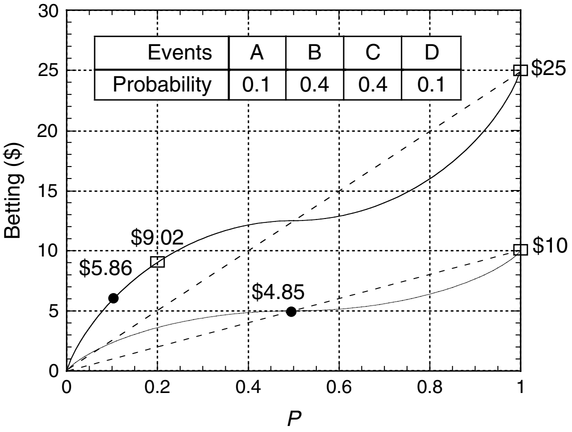
Let P_i and V_i be the probability that the event-i occurs and the utility function of event-i, respectively. People act as the expected utility that is expressed as

$$\sum_i P_i V_i$$

takes the maximum value.

5. Noteworthy point

(a) Almost all data reported by Tversky and Fox[1] can be sufficiently explained by the degree of anxiety/expectation.



Events	A	B	C	D
Probability	0.1	0.4	0.4	0.1

Figure 6.9 Expectation–probability curve of betting for certain condition that gives reasonable explanation.

(b) A newly defined degree of anxiety/expectation is very useful in the field of human engineering.

(c) In Figure 6.9, an example (the probability of occurrence of each event is set as $P(\mathrm{A}) = 0.1$, $P(\mathrm{B}) = 0.4$, $P(\mathrm{C}) = 0.4$, and $P(\mathrm{D}) = 0.1$) that satisfies all the above choices is shown. In other words, all the patterns described in the table are explained by the new degree of anxiety/expectation in Eq. (6.7). When the relation between the probability of occurrence and the degree of expectation is linear, it is impossible to use the above discussion.

6.5 Decision-making regarding daily insignificant matters

When decision-making is done with a steady interest, as described above, the weight of judgment for each probability of occurrence should be set to a constant value. However, people have a tendency to judge matters according to an intuitive and simple consideration process without a strict consideration process based on the probability terms. When a daily trifling matter is considered for decision-making without sufficient interest or the decision-making is done in an instant without sufficient consideration, the weight of judgment for the small value of the probability of occurrence usually becomes smaller than that for the high value of the probability of occurrence. Therefore, the degree of anxiety/expectation in such cases should be expressed as a multiplication of Eq. (6.7) and the function of the weight of judgment. The function of the weight of judgment should be zero at $P = 0$ and unity at $P = 1$. As a simple expression of the function of the weight of judgment, a power function of the form P^n can be approximated. This form satisfies the necessary conditions described above. When P^n is adopted as the function of the weight of judgment, the new degree of anxiety/expectation can be expressed as follows:

$$\mathrm{AE}_{P<1/2} = P^n V\{-P\ln P - (1-P)\ln(1-P)\} \tag{6.8a}$$

$$\mathrm{AE}_{P<1/2} = P^n V[2\ln 2 - \{-P\ln P - (1-P)\ln(1-P)\}] \tag{6.8b}$$

Figure 6.10 shows the relationship between P and P^n. When $n = 0$, the function of the weight of judgment becomes unity regardless of the value of probability of occurrence; this case corresponds to the case where the weight of judgment for each probability of occurrence is set to a constant value. There is no need to consider the function of the weight of judgment as described above. By changing the value of n, the degree of anxiety/expectation under the condition of $V = 1$ is calculated and some example of the results are shown in Figure 6.11. It is clarified that the curve of the degree of anxiety/expectation depends on the value of n, and the curve becomes monotonous in proportion to an increase in the value of n.

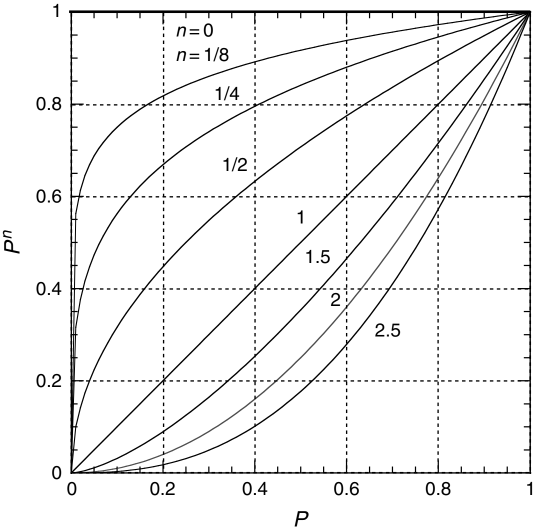


Figure 6.10 Distributions of weight function.

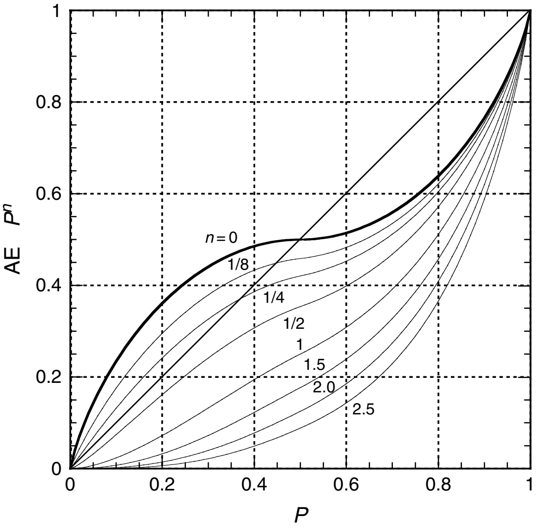


Figure 6.11 Anxiety–probability curves considering weight function.

When the function of the weight of judgment is considered, the critical value of the probability to make a decision has higher values than that without considering the function of the weight of judgment. The result at approximately $n = 1/4$ is close to the data in the paper by Tversky and Fox[1]; in other words, the degree of anxiety/expectation overweighs small probabilities and underweighs not only high probabilities but also moderate probabilities. However, it is still necessary to consider the function of the weight of judgment when the decision-making is performed with a steady interest.

6.6 Summary

In this chapter, the following points have been clarified:

(1) The degree of anxiety/expectation can be defined by making use of information entropy. This index has a clear physical background.
(2) It is possible to make decisions based on the newly defined index.
(3) By making use of the newly defined index, many results that are impossible to explain by using the old line of linear thought can be explained.
(4) The usefulness of the newly defined index is clarified by applying it to the following examples and experimental results:
    (a) anxiety regarding possible involvement in an accident when going outdoors
    (b) priority for improvement
    (c) Should the proposal for improvement for increase in the gain be adopted?
    (d) Should the proposal for improvement for decrease in the loss be adopted?
    (e) collation with earlier discussions by Tversky and Fox[1]

References

Chapter 1

1. Abramson, N. (1963). *Information Theory and Coding*. McGraw-Hill.
2. Brillouin, L. (1962). *Science and Information Theory*. Academic Press.
3. Segawa, F. (1972). Entropy in physical phenomena. *Math. Sci.*, **8**, 11–15.
4. Shanon, C.E. (1949). *The Mathematical Theory of Communication*. The University of Illinois Press.

Chapter 2

1. Ito, S. and Ogawa, K. (1978). Evaluation of mixing and separation capacity and information entropy. *Kagaku Kogaku*, **42**, 210–215.
2. Ito, S. and Ogawa, K. (1975). A definition of quality of mixedness. *J. Chem. Eng. Jpn*, **8**, 148–151.
3. Ito, S., Ogawa, K., and Matsumura, Y. (1980). Mixing rate in a stirred vessel. *J. Chem. Eng. Jpn*, **13**, 324–326.
4. Laine, J. (1983). Ruhrintensitat und Leistung von Scheiben und Lochscheibenruhren im groβtechnishen Maβtab. *Chem. Ing. Tech.*, **55**, 574–575.
5. Ogawa, K. and Ito, S. (1974). Turbulent mixing phenomena in a circular pipe. *Kagaku Kogaku*, **38**, 815–819.
6. Lawn, C.J. (1971). The determination of the rate of dissipation in turbulent pipe flow. *J. Fluid Mech.*, **48**, 477–505.
7. Laufer, J. (1954). The structure of turbulence in fully developed pipe flow. NACA TN, 1–16.
8. Quarmby, A. and Anand, R.K. (1969). Axisymmetric turbulent mass transfer in circular pipe tube. *J. Fluid Mech.*, **38**, 433–455.
9. Ogawa, K. (1981). Definition of local and whole mixing capacity indexes of equipment. *Kagaku Kogaku Ronbunshu*, **7**, 207–210.
10. Ogawa, K. and Kuroda, C. (1984). Local mixing capacity of turbulent flow in a circular pipe. *Kagaku Kogaku Ronbunshu*, **10**, 268–271.
11. Ogawa, K. (1984). An expression of quality of mixedness for multi-component batch mixing. *Kagaku Kogaku Ronbunshu*, **10**, 261–264.
12. Ogawa, K. (1992). Saikin no Kagaku Kogaku, Kagaku-Kogyo-Sha.

Chapter 3

1. Ogawa, K., Ito, S., and Kishino, H. (1978). A definition of separation efficiency. *J. Chem. Eng. Jpn*, **11**, 44–47.

Chapter 4

1. Chandrasekhar, S. (1949). On Heisenberg's elementary theory of turbulence. *Proc. Rpy. Soc. Lond. Ser. A*, **200**, 20–33.
2. Rotta, J. (1950). Das Spektrum Isotroper Turbulenz in Statstishen Gleichgewicht. *Ingr. Arch.*, **18**, 60–76.
3. Prudman, I. (1951). A comparison of Heisenberg's spectrum of turbulence with experiment. *Proc. Camb. Phil. Soc.*, **47**, 158–176.
4. Birkhoff, G. (1954). Fourier synthesis of homogeneous turbulence. *Commun. Pure Appl. Math.*, **7**, 19–44.
5. Loitsansky, L.G. (1939). Some basic laws of isotropic turbulent flow. NACA, TM-1079.
6. Ogawa, K. (1981). A simple formula of energy spectrum function in low wavenumber ranges. *J. Chem. Eng. Jpn*, **14**, 250–252.
7. Kolmogoroff, A.N. (1941). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *C. R. Acad. URSS*, **30**, 301.
8. Heisenberg, W. (1948). Zur Statistischen Theorie der Turbulenz. *Z. Phys.*, **124**, 628–657.
9. Ogawa, K., Kuroda, C., and Yoshikawa, S. (1985). An expression of energy spectrum function for wide wavenumber ranges. *J. Chem. Eng. Jpn*, **18**, 544–549.
10. Ogawa, K. (1988). Energy spectrum function of eddy group gathered together as a model of turbulence. *Int. J. Eng. Fluid Mech.*, **1**, 235–244.
11. Ogawa, K. (1993). Non-linear fantasy of fluid flow-order of turbulence as chaos and scale-up. *Kagaku Kogaku*, **57**, 120–123.
12. Ogawa, K., Kuroda, C., and Yoshikawa, S. (1986). A method of scaling up equipment from the viewpoint of energy spectrum function. *J. Chem. Eng. Jpn*, **19**, 345–347.
13. Ogawa, K. (1992). Evaluation of common scaling-up rules for a stirred vessel from the viewpoint of energy spectrum function. *J. Chem. Eng. Jpn*, **25**, 750–752.
14. Gibson, C.H. and Schwarz, W.H. (1963). The universal equilibrium spectra of turbulent velocity and scalar field. *J. Fluid Mech.*, **16**, 365–384.
15. Laufer, J. (1954). The structure of turbulence in fully developed pipe flow. *NACA*, TN1174, 1–16.
16. Lawn, J. (1971). The determination of the rate of dissipation in turbulent pipe flow, *J. Fluid Mech.*, **48**, 477–505.
17. Stewart, R.W. and Townsend, A.A. (1951). Similarity and self-preservation in isotropic turbulence. *Phil. Trans. Roy. Soc. London*, **243A**, 359–386.
18. Ogawa, K. (2000). Kagaku Kogaku no Shinp034-Mikishinngu gijutu, Kagaku Kogakkai, Maki-shoten.
19. Ogawa, K. (2001). *Supekutoru Mitudo Kannsuu to sono Ouyou*. Asakura-shoten.

Chapter 5

1. Ogawa, K. (1986). Effectiveness of information entropy for evaluation of grinding efficiency. *Chem. Eng. Commun.*, **46**, 1–9.
2. Ogawa, K. (1990). A single expression of common particle size distribution. *Part. Part. Syst. Charact.*, **7**, 127–130.
3. Ogawa, K. (2000). Kagaku Kogaku no Shinpo34-Mikishinngu gijutu, Kagaku Kogakkai, Maki-shoten.

Chapter 6

1. Tversky, A. and Fox, C.R. (1995). Weighing risk and uncertainty. *Psychol. Rev.*, **102**, 269–283.
2. Ogawa, K. (2006). Quantitative index for anxiety/expectation and its applications. *J. Chem. Eng. Jpn*, **39**, 102–110.

This page intentionally left blank

Epilogue

The author intended to show the usefulness of information entropy in chemical engineering although information entropy is not common in this field. Not many examples have been considered for explaining the usefulness of information entropy. Therefore, it is still a question whether information entropy can be applied in chemical engineering. However, the author expects that information entropy will be accepted in chemical engineering, since it is related to the probability terms that are widely used in chemical engineering. Since there are many phenomena that can be expressed by probability terms in chemical engineering, information entropy would be valid over a wide area. The author expects that the number of researchers interested in information entropy will increase.

This page intentionally left blank

Index